



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification 7 :</b> <b>C12N 15/54, 1/21, 9/10, C12Q 1/48, 1/68</b>		<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 00/04166</b>
			<b>(43) International Publication Date:</b> 27 January 2000 (27.01.00)
<b>(21) International Application Number:</b> PCT/US99/15871		<b>(74) Agent:</b> MAJARIAN, William, R.; E.I. du Pont de Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).	
<b>(22) International Filing Date:</b> 13 July 1999 (13.07.99)			
<b>(30) Priority Data:</b> 60/092,844 14 July 1998 (14.07.98) US		<b>(81) Designated States:</b> AE, AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, US, UZ, VN, YU, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
<b>(71) Applicant (for all designated States except US):</b> E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).			
<b>(72) Inventors; and</b>			
<b>(75) Inventors/Applicants (for US only):</b> ALLEN, Stephen, M. [US/US]; 2225 Rosewood Drive, Wilmington, DE 19810 (US). FADER, Gary, M. [US/US]; 1000 Woods Lane, Landenberg, PA 19350 (US). FALCO, Saverio, Carl [US/US]; 1902 Millers Road, Arden, DE 19810 (US). KINNEY, Anthony, J. [GB/US]; 609 Lore Avenue, Wilmington, DE 19809 (US). LIGHTNER, Jonathan, E. [US/US]; 4180 Delta Road, Airville, PA 17302 (US). MIAO, Guo-Hua [CN/US]; 202 Cherry Blossom Place, Hockessin, DE 19707 (US). RAFALSKI, J., Antoni [US/US]; 2028 Longcome Drive, Wilmington, DE 19810 (US). THORPE, Catherine, J. [GB/GB]; 120 Ross Street, Cambridge CB1 3BU (GB).			
<b>(54) Title:</b> PLANT CELLULOSE SYNTHASES			
<b>(57) Abstract</b>			
<p>This invention relates to an isolated nucleic acid fragment encoding a cellulose synthase. The invention also relates to the construction of a chimeric gene encoding all or a portion of the cellulose synthase, in sense or antisense orientation, wherein expression of the chimeric gene results in production of altered levels of the cellulose synthase in a transformed host cell.</p>			
<pre>SEQ ID NO:1 1 60 RANQAGHREKQPEQKIASVLP--LPESATPTTPPRKRYR--RTACPG--I SEQ ID NO:4 SEQ ID NO:6 SEQ ID NO:8 SEQ ID NO:10 SEQ ID NO:12 SEQ ID NO:14 SEQ ID NO:16 SEQ ID NO:18 SEQ ID NO:20 SEQ ID NO:22 SEQ ID NO:24 SEQ ID NO:26 SEQ ID NO:28 SEQ ID NO:29 R--FR--  SEQ ID NO:2 61 120 W-ESQARG--ME-ASAGLVAGSHHSELY-VIRLDGEPQKE--MDQAGQVCI-- SEQ ID NO:4 SEQ ID NO:6 SEQ ID NO:8 SEQ ID NO:10 SEQ ID NO:12 SEQ ID NO:14 SEQ ID NO:16 SEQ ID NO:18 SEQ ID NO:20 SEQ ID NO:22 SEQ ID NO:24 SEQ ID NO:26 SEQ ID NO:28 SEQ ID NO:29 L-AGSINREFF--LI--HDESHARISVQELSGQTCQI--  SEQ ID NO:2 121 180 CGDGVGMPDGEFFVACRECAFFICDCTETEDKRGVQPCQCTRYKLLGCMVFGD- SEQ ID NO:4 SEQ ID NO:6 SEQ ID NO:8 SEQ ID NO:10 SEQ ID NO:12 SEQ ID NO:14 SEQ ID NO:16 SEQ ID NO:18 SEQ ID NO:20 SEQ ID NO:22 SEQ ID NO:24 SEQ ID NO:26 SEQ ID NO:28 SEQ ID NO:29 CGDGVGMPDGEFFVACRECAFFICDCTETEDKRGVQPCQCTRYKLLGCMVFGD-</pre>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LJ	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

TITLE

## PLANT CELLULOSE SYNTHASES

This application claims the benefit of U.S. Provisional Application No. 60/092,844, filed July 14, 1998.

FIELD OF THE INVENTION

This invention is in the field of plant molecular biology. More specifically, this invention pertains to nucleic acid fragments encoding cellulose biosynthetic enzymes in plants and seeds.

BACKGROUND OF THE INVENTION

Cellulose is a major component of plant fiber, e.g. cotton fiber. Cellulose is composed of crystalline beta-1,4-glucan microfibrils (see World Patent Publication No. WO 98/00549). These microfibrils are strong and can resist enzymatic and mechanical degradation and are important in determining nutritional quality of animal and human foodstuffs. Hence, modification of the biosynthetic pathway responsible for cellulose synthesis through modification of cellulose synthase activity could potentially alter fiber quantity, either by producing more or less fiber in a particular plant species or in a specific organ or tissue of a particular plant. Modification of cellulose synthase activity could increase the value of the fiber to the end-user and may improve the structural integrity of the plant cell wall. Lastly, because cellulose is a major cell wall component, inhibition of cellulose synthesis would probably be lethal. Thus, cellulose synthase may serve as the target for a novel class of herbicides. Plant cellulose synthase genes, homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase, have been reported from cotton, *Arabidopsis*, rice and alfalfa (World Patent Publication Nos. WO 98/00549 and WO 98/18949).

There is a great deal of interest in identifying the genes that encode proteins involved in cellulose synthesis. These genes may be used in plant cells to control the synthesis of cellulose. Accordingly, the availability of nucleic acid sequences encoding all or a portion of a cellulose synthase would facilitate studies to better understand cellulose synthesis in plants and provide genetic tools to alter cellulose production.

SUMMARY OF THE INVENTION

The instant invention relates to isolated nucleic acid fragments encoding cellulose biosynthesis enzymes. Specifically, this invention concerns an isolated nucleic acid fragment encoding a cellulose synthase and an isolated nucleic acid fragment that is substantially similar to an isolated nucleic acid fragment encoding a cellulose synthase. In addition, this invention relates to a nucleic acid fragment that is complementary to the nucleic acid fragment encoding cellulose synthase. An additional embodiment of the instant invention pertains to a polypeptide encoding all or a substantial portion of a cellulose synthase.

In another embodiment, the instant invention relates to a chimeric gene encoding a cellulose synthase, or to a chimeric gene that comprises a nucleic acid fragment that is complementary to a nucleic acid fragment encoding a cellulose synthase, operably linked to suitable regulatory sequences, wherein expression of the chimeric gene results in production of levels of the encoded protein in a transformed host cell that is altered (i.e., increased or decreased) from the level produced in an untransformed host cell.

In a further embodiment, the instant invention concerns a transformed host cell comprising in its genome a chimeric gene encoding a cellulose synthase, operably linked to suitable regulatory sequences. Expression of the chimeric gene results in production of altered levels of the encoded protein in the transformed host cell. The transformed host cell can be of eukaryotic or prokaryotic origin, and include cells derived from higher plants and microorganisms. The invention also includes transformed plants that arise from transformed host cells of higher plants, and seeds derived from such transformed plants.

An additional embodiment of the instant invention concerns a method of altering the level of expression of a cellulose synthase in a transformed host cell comprising:  
a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding a cellulose synthase; and b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of altered levels of cellulose synthase in the transformed host cell.

An addition embodiment of the instant invention concerns a method for obtaining a nucleic acid fragment encoding all or a substantial portion of an amino acid sequence encoding a cellulose synthase.

A further embodiment of the instant invention is a method for evaluating at least one compound for its ability to inhibit the activity of a cellulose synthase, the method comprising the steps of: (a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding a cellulose synthase, operably linked to suitable regulatory sequences; (b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of cellulose synthase in the transformed host cell; (c) optionally purifying the cellulose synthase expressed by the transformed host cell; (d) treating the cellulose synthase with a compound to be tested; and (e) comparing the activity of the cellulose synthase that has been treated with a test compound to the activity of an untreated cellulose synthase, thereby selecting compounds with potential for inhibitory activity.

#### BRIEF DESCRIPTION OF THE DRAWINGS AND SEQUENCE DESCRIPTIONS

The invention can be more fully understood from the following detailed description and the accompanying drawings and Sequence Listing which form a part of this application.



Figure 1 shows a comparison of the amino acid sequences set forth in SEQ ID NOs:2, 4, 8, 10, 12, 14, 16, 18, 20 and 22 and the *Arabidopsis thaliana* sequences (SEQ ID NOs:23 (gi 2827139), 24 (gi 2827141), 26 (gi 4467125), 27 (gi 4886756) and 29 (gi 3135611)) and *Gossypium hirsutum* sequences (SEQ ID NOs:25 (gi 1706958) and 28 (gi 5081779)).

- 5 Table 1 lists the polypeptides that are described herein, the designation of the cDNA clones that comprise the nucleic acid fragments encoding polypeptides representing all or a substantial portion of these polypeptides, and the corresponding identifier (SEQ ID NO:) as used in the attached Sequence Listing. The sequence descriptions and Sequence Listing attached hereto comply with the rules governing nucleotide and/or amino acid sequence  
10 disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

**TABLE 1**  
Cellulose Biosynthetic Enzymes

Protein	Clone Designation	SEQ ID NO:	
		(Nucleotide)	(Amino Acid)
Cellulose Synthase	bsh1.pk0002.f6	1	2
Cellulose Synthase	Contig composed of: cco1n.pk0005.g3 cdt2c.pk002.g1 cdt2c.pk002.l16 csc1c.pk002.i1 p0031.ccmr05rb p0110.cgsm57r	3	4
Cellulose Synthase	cr1n.pk0135.e10	5	6
Cellulose Synthase	p0097.cqrad17rc	7	8
Cellulose Synthase	p0122.ckamh70rc	9	10
Cellulose Synthase	rlr24.pk0073.g1	11	12
Cellulose Synthase	sdp2c.pk005.o22	13	14
Cellulose Synthase	ses8w.pk0028.f3	15	16
Cellulose Synthase	ssl.pk0036.c10	17	18
Cellulose Synthase	Contig composed of: wl1.pk0009.c9 wr1.pk0160.d11 wre1n.pk0043.f9 wre1n.pk0043.h8 wre1n.pk0131.g10	19	20
Cellulose Synthase	wl1n.pk0044.b1	21	22

- 15 The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Research* 13:3021-3030 (1985) and in the *Biochemical Journal* 219 (No. 2):345-373 (1984) which are herein incorporated by reference. The

symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

#### DETAILED DESCRIPTION OF THE INVENTION

In the context of this disclosure, a number of terms shall be utilized. As used herein, a  
5 “nucleic acid fragment” is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. A nucleic acid fragment in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

As used herein, “contig” refers to a nucleotide sequence that is assembled from two or  
10 more constituent nucleotide sequences that share common or overlapping regions of sequence homology. For example, the nucleotide sequences of two or more nucleic acid fragments can be compared and aligned in order to identify common or overlapping sequences. Where common or overlapping sequences exist between two or more nucleic acid fragments, the sequences (and thus their corresponding nucleic acid fragments) can be  
15 assembled into a single contiguous nucleotide sequence.

As used herein, “substantially similar” refers to nucleic acid fragments wherein changes in one or more nucleotide bases results in substitution of one or more amino acids, but do not affect the functional properties of the polypeptide encoded by the nucleotide sequence. “Substantially similar” also refers to nucleic acid fragments wherein changes in  
20 one or more nucleotide bases does not affect the ability of the nucleic acid fragment to mediate alteration of gene expression by gene silencing through for example antisense or co-suppression technology. “Substantially similar” also refers to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotides that do not substantially affect the functional properties of the resulting  
25 transcript vis-à-vis the ability to mediate gene silencing or alteration of the functional properties of the resulting protein molecule. It is therefore understood that the invention encompasses more than the specific exemplary nucleotide or amino acid sequences and includes functional equivalents thereof.

For example, it is well known in the art that antisense suppression and co-suppression  
30 of gene expression may be accomplished using nucleic acid fragments representing less than the entire coding region of a gene, and by nucleic acid fragments that do not share 100% sequence identity with the gene to be suppressed. Moreover, alterations in a nucleic acid fragment which result in the production of a chemically equivalent amino acid at a given site, but do not effect the functional properties of the encoded polypeptide, are well known in  
35 the art. Thus, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in substitution of one negatively charged residue for another, such as aspartic acid for

glutamic acid, or one positively charged residue for another, such as lysine for arginine, can also be expected to produce a functionally equivalent product. Nucleotide changes which result in alteration of the N-terminal and C-terminal portions of the polypeptide molecule would also not be expected to alter the activity of the polypeptide. Each of the proposed  
5 modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products.

Moreover, substantially similar nucleic acid fragments may also be characterized by their ability to hybridize, under stringent conditions (0.1X SSC, 0.1% SDS, 65°C), with the nucleic acid fragments disclosed herein.

10 Substantially similar nucleic acid fragments of the instant invention may also be characterized by the percent identity of the amino acid sequences that they encode to the amino acid sequences disclosed herein, as determined by algorithms commonly employed by those skilled in this art. Preferred are those nucleic acid fragments whose nucleotide sequences encode amino acid sequences that are 80% identical to the amino acid sequences  
15 reported herein. More preferred nucleic acid fragments encode amino acid sequences that are 90% identical to the amino acid sequences reported herein. Most preferred are nucleic acid fragments that encode amino acid sequences that are 95% identical to the amino acid sequences reported herein. Sequence alignments and percent identity calculations were performed using the Megalign program of the LASARGENE bioinformatics computing suite  
20 (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS*. 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

25 A "substantial portion" of an amino acid or nucleotide sequence comprises an amino acid or a nucleotide sequence that is sufficient to afford putative identification of the protein or gene that the amino acid or nucleotide sequence comprises. Amino acid and nucleotide sequences can be evaluated either manually by one skilled in the art, or by using computer-based sequence comparison and identification tools that employ algorithms such as BLAST  
30 (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol.* 215:403-410; see also [www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)). In general, a sequence of ten or more contiguous amino acids or thirty or more contiguous nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene-specific oligonucleotide probes  
35 comprising 30 or more contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12 or more nucleotides may be used as amplification primers in PCR in order to obtain a particular

nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises a nucleotide sequence that will afford specific identification and/or isolation of a nucleic acid fragment comprising the sequence. The instant specification teaches amino acid and nucleotide sequences encoding polypeptides that  
5 comprise one or more particular plant proteins. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

10 "Codon degeneracy" refers to divergence in the genetic code permitting variation of the nucleotide sequence without effecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleic acid fragment comprising a nucleotide sequence that encodes all or a substantial portion of the amino acid sequences set forth herein. The skilled artisan is well aware of the "codon-bias" exhibited  
15 by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a nucleic acid fragment for improved expression in a host cell, it is desirable to design the nucleic acid fragment such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

"Synthetic nucleic acid fragments" can be assembled from oligonucleotide building  
20 blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form larger nucleic acid fragments which may then be enzymatically assembled to construct the entire desired nucleic acid fragment. "Chemically synthesized", as related to nucleic acid fragment, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of nucleic acid fragments  
25 may be accomplished using well established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines. Accordingly, the nucleic acid fragments can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards  
30 those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell where sequence information is available.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its  
35 own regulatory sequences. "Chimeric gene" refers any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived

from the same source, but arranged in a manner different than that found in nature.

“Endogenous gene” refers to a native gene in its natural location in the genome of an organism. A “foreign” gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A “transgene” is a gene that has been introduced into the genome by a transformation procedure.

“Coding sequence” refers to a nucleotide sequence that codes for a specific amino acid sequence. “Regulatory sequences” refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, introns, and polyadenylation recognition sequences.

“Promoter” refers to a nucleotide sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. The promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an “enhancer” is a nucleotide sequence which can stimulate promoter activity and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic nucleotide segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. Promoters which cause a nucleic acid fragment to be expressed in most cell types at most times are commonly referred to as “constitutive promoters”. New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro and Goldberg (1989) *Biochemistry of Plants* 15:1-82. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, nucleic acid fragments of different lengths may have identical promoter activity.

The “translation leader sequence” refers to a nucleotide sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner and Foster (1995) *Molecular Biotechnology* 3:225).

The “3' non-coding sequences” refer to nucleotide sequences located downstream of a coding sequence and include polyadenylation recognition sequences and other sequences

encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The use of different 3' non-coding sequences is exemplified by Ingelbrecht et al. (1989) *Plant Cell* 1:671-680.

5       “RNA transcript” refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. “Messenger RNA (mRNA)” refers to the RNA that is  
10 without introns and that can be translated into polypeptide by the cell. “cDNA” refers to a double-stranded DNA that is complementary to and derived from mRNA. “Sense” RNA refers to an RNA transcript that includes the mRNA and so can be translated into a polypeptide by the cell. “Antisense RNA” refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the  
15 expression of a target gene (see U.S. Patent No. 5,107,065, incorporated herein by reference). The complementarity of an antisense RNA may be with any part of the specific nucleotide sequence, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. “Functional RNA” refers to sense RNA, antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes.  
20       The term “operably linked” refers to the association of two or more nucleic acid fragments on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to  
25 regulatory sequences in sense or antisense orientation.

      The term “expression”, as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide. “Antisense inhibition” refers to the production of antisense RNA transcripts capable of  
30 suppressing the expression of the target protein. “Overexpression” refers to the production of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms. “Co-suppression” refers to the production of sense RNA transcripts capable of suppressing the expression of identical or substantially similar foreign or endogenous genes (U.S. Patent No. 5,231,020, incorporated herein by reference).

35       “Altered levels” refers to the production of gene product(s) in transgenic organisms in amounts or proportions that differ from that of normal or non-transformed organisms.

      “Mature” protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or propeptides present in the primary translation product have been removed.

"Precursor" protein refers to the primary product of translation of mRNA; i.e., with pre- and propeptides still present. Pre- and propeptides may be but are not limited to intracellular localization signals.

5 A "chloroplast transit peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the chloroplast or other plastid types present in the cell in which the protein is made. "Chloroplast transit sequence" refers to a nucleotide sequence that encodes a chloroplast transit peptide. A "signal peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the secretory system (Chrispeels (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol.* 42:21-53).  
10 If the protein is to be directed to a vacuole, a vacuolar targeting signal (*supra*) can further be added, or if to the endoplasmic reticulum, an endoplasmic reticulum retention signal (*supra*) may be added. If the protein is to be directed to the nucleus, any signal peptide present should be removed and instead a nuclear localization signal included (Raikhel (1992) *Plant Phys.* 100:1627-1632).  
15 "Transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" organisms. Examples of methods of plant transformation include *Agrobacterium*-mediated transformation (De Blaere et al. (1987) *Meth. Enzymol.* 143:277) and particle-accelerated or "gene gun" transformation  
20 technology (Klein et al. (1987) *Nature (London)* 327:70-73; U.S. Patent No. 4,945,050, incorporated herein by reference).

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook et al. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989  
25 (hereinafter "Maniatis").

Nucleic acid fragments encoding at least a portion of a cellulose synthase enzyme have been isolated and identified by comparison of random plant cDNA sequences to public databases containing nucleotide and protein sequences using the BLAST algorithms well known to those skilled in the art. The nucleic acid fragments of the instant invention may be  
30 used to isolate cDNAs and genes encoding homologous proteins from the same or other plant species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to, methods of nucleic acid hybridization, and methods of DNA and RNA amplification as exemplified by various uses of nucleic acid amplification technologies (e.g., polymerase  
35 chain reaction, ligase chain reaction).

For example, genes encoding other cellulose synthase enzymes, either as cDNAs or genomic DNAs, could be isolated directly by using all or a portion of the instant nucleic acid fragments as DNA hybridization probes to screen libraries from any desired plant employing

methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the instant nucleic acid sequences can be designed and synthesized by methods known in the art (Maniatis). Moreover, the entire sequences can be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primer DNA labeling, nick translation, or end-labeling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers can be designed and used to amplify a part or all of the instant sequences. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as probes to isolate full length cDNA or genomic fragments under conditions of appropriate stringency.

In addition, two short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols to amplify longer nucleic acid fragments encoding homologous genes from DNA or RNA. The polymerase chain reaction may also be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the instant nucleic acid fragments, and the sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, the skilled artisan can follow the RACE protocol (Frohman et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:8998) to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Primers oriented in the 3' and 5' directions can be designed from the instant sequences. Using commercially available 3' RACE or 5' RACE systems (BRL), specific 3' or 5' cDNA fragments can be isolated (Ohara et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:5673; Loh et al. (1989) *Science* 243:217). Products generated by the 3' and 5' RACE procedures can be combined to generate full-length cDNAs (Frohman and Martin (1989) *Techniques* 1:165).

Availability of the instant nucleotide and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing portions of the instant amino acid sequences may be synthesized. These peptides can be used to immunize animals to produce polyclonal or monoclonal antibodies with specificity for peptides or proteins comprising the amino acid sequences. These antibodies can be then be used to screen cDNA expression libraries to isolate full-length cDNA clones of interest (Lerner (1984) *Adv. Immunol.* 36:1; Maniatis).

The nucleic acid fragments of the instant invention may be used to create transgenic plants in which the disclosed polypeptides are present at higher or lower levels than normal or in cell types or developmental stages in which they are not normally found. This would have the effect of altering the level of cellulose synthase in those cells.



Overexpression of the proteins of the instant invention may be accomplished by first constructing a chimeric gene in which the coding region is operably linked to a promoter capable of directing expression of a gene in the desired tissues at the desired stage of development. For reasons of convenience, the chimeric gene may comprise promoter sequences and translation leader sequences derived from the same genes. 3' Non-coding sequences encoding transcription termination signals may also be provided. The instant chimeric gene may also comprise one or more introns in order to facilitate gene expression.

Plasmid vectors comprising the instant chimeric gene can then be constructed. The choice of plasmid vector is dependent upon the method that will be used to transform host plants. The skilled artisan is well aware of the genetic elements that must be present on the plasmid vector in order to successfully transform, select and propagate host cells containing the chimeric gene. The skilled artisan will also recognize that different independent transformation events will result in different levels and patterns of expression (Jones et al. (1985) *EMBO J.* 4:2411-2418; De Almeida et al. (1989) *Mol. Gen. Genetics* 218:78-86), and thus that multiple events must be screened in order to obtain lines displaying the desired expression level and pattern. Such screening may be accomplished by Southern analysis of DNA, Northern analysis of mRNA expression, Western analysis of protein expression, or phenotypic analysis.

For some applications it may be useful to direct the instant polypeptides to different cellular compartments, or to facilitate its secretion from the cell. It is thus envisioned that the chimeric gene described above may be further supplemented by altering the coding sequence to encode the instant polypeptides with appropriate intracellular targeting sequences such as transit sequences (Keegstra (1989) *Cell* 56:247-253), signal sequences or sequences encoding endoplasmic reticulum localization (Chrispeels (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol.* 42:21-53), or nuclear localization signals (Raikhel (1992) *Plant Phys.* 100:1627-1632) added and/or with targeting sequences that are already present removed. While the references cited give examples of each of these, the list is not exhaustive and more targeting signals of utility may be discovered in the future.

It may also be desirable to reduce or eliminate expression of genes encoding the instant polypeptides in plants for some applications. In order to accomplish this, a chimeric gene designed for co-suppression of the instant polypeptide can be constructed by linking a gene or gene fragment encoding that polypeptide to plant promoter sequences. Alternatively, a chimeric gene designed to express antisense RNA for all or part of the instant nucleic acid fragment can be constructed by linking the gene or gene fragment in reverse orientation to plant promoter sequences. Either the co-suppression or antisense chimeric genes could be introduced into plants via transformation wherein expression of the corresponding endogenous genes are reduced or eliminated.

Molecular genetic solutions to the generation of plants with altered gene expression have a decided advantage over more traditional plant breeding approaches. Changes in plant phenotypes can be produced by specifically inhibiting expression of one or more genes by antisense inhibition or cosuppression (U. S. Patent Nos. 5,190,931, 5,107,065 and 5,283,323). An antisense or cosuppression construct would act as a dominant negative regulator of gene activity. While conventional mutations can yield negative regulation of gene activity these effects are most likely recessive. The dominant negative regulation available with a transgenic approach may be advantageous from a breeding perspective. In addition, the ability to restrict the expression of specific phenotype to the reproductive tissues of the plant by the use of tissue specific promoters may confer agronomic advantages relative to conventional mutations which may have an effect in all tissues in which a mutant gene is ordinarily expressed.

The person skilled in the art will know that special considerations are associated with the use of antisense or cosuppression technologies in order to reduce expression of particular genes. For example, the proper level of expression of sense or antisense genes may require the use of different chimeric genes utilizing different regulatory elements known to the skilled artisan. Once transgenic plants are obtained by one of the methods described above, it will be necessary to screen individual transgenics for those that most effectively display the desired phenotype. Accordingly, the skilled artisan will develop methods for screening large numbers of transformants. The nature of these screens will generally be chosen on practical grounds, and is not an inherent part of the invention. For example, one can screen by looking for changes in gene expression by using antibodies specific for the protein encoded by the gene being suppressed, or one could establish assays that specifically measure enzyme activity. A preferred method will be one which allows large numbers of samples to be processed rapidly, since it will be expected that a large number of transformants will be negative for the desired phenotype.

The instant polypeptides (or portions thereof) may be produced in heterologous host cells, particularly in the cells of microbial hosts, and can be used to prepare antibodies to these proteins by methods well known to those skilled in the art. The antibodies are useful for detecting the polypeptides of the instant invention *in situ* in cells or *in vitro* in cell extracts. Preferred heterologous host cells for production of the instant polypeptides are microbial hosts. Microbial expression systems and expression vectors containing regulatory sequences that direct high level expression of foreign proteins are well known to those skilled in the art. Any of these could be used to construct a chimeric gene for production of the instant polypeptides. This chimeric gene could then be introduced into appropriate microorganisms via transformation to provide high level expression of the encoded cellulose synthase. An example of a vector for high level expression of the instant polypeptides in a bacterial host is provided (Example 6).

Additionally, the instant polypeptides can be used as a targets to facilitate design and/or identification of inhibitors of those enzymes that may be useful as herbicides. This is desirable because the polypeptides described herein catalyze a step in the synthesis of cellulose. Accordingly, inhibition of the activity of one or more of the enzymes described  
5 herein could lead to inhibition plant growth. Thus, the instant polypeptides could be appropriate for new herbicide discovery and design.

All or a substantial portion of the nucleic acid fragments of the instant invention may also be used as probes for genetically and physically mapping the genes that they are a part of, and as markers for traits linked to those genes. Such information may be useful in plant  
10 breeding in order to develop lines with desired phenotypes. For example, the instant nucleic acid fragments may be used as restriction fragment length polymorphism (RFLP) markers. Southern blots (Maniatis) of restriction-digested plant genomic DNA may be probed with the nucleic acid fragments of the instant invention. The resulting banding patterns may then be subjected to genetic analyses using computer programs such as MapMaker (Lander et al.  
15 (1987) *Genomics* 1:174-181) in order to construct a genetic map. In addition, the nucleic acid fragments of the instant invention may be used to probe Southern blots containing restriction endonuclease-treated genomic DNAs of a set of individuals representing parent and progeny of a defined genetic cross. Segregation of the DNA polymorphisms is noted and used to calculate the position of the instant nucleic acid sequence in the genetic map  
20 previously obtained using this population (Botstein et al. (1980) *Am. J. Hum. Genet.* 32:314-331).

The production and use of plant gene-derived probes for use in genetic mapping is described in Bernatzky and Tanksley (1986) *Plant Mol. Biol. Reporter* 4(1):37-41. Numerous publications describe genetic mapping of specific cDNA clones using the  
25 methodology outlined above or variations thereof. For example, F2 intercross populations, backcross populations, randomly mated populations, near isogenic lines, and other sets of individuals may be used for mapping. Such methodologies are well known to those skilled in the art.

Nucleic acid probes derived from the instant nucleic acid sequences may also be used  
30 for physical mapping (i.e., placement of sequences on physical maps; see Hoheisel et al. In: *Nonmammalian Genomic Analysis: A Practical Guide*, Academic press 1996, pp. 319-346, and references cited therein).

In another embodiment, nucleic acid probes derived from the instant nucleic acid sequences may be used in direct fluorescence *in situ* hybridization (FISH) mapping (Trask  
35 (1991) *Trends Genet.* 7:149-154). Although current methods of FISH mapping favor use of large clones (several to several hundred KB; see Laan et al. (1995) *Genome Research* 5:13-20), improvements in sensitivity may allow performance of FISH mapping using shorter probes.

A variety of nucleic acid amplification-based methods of genetic and physical mapping may be carried out using the instant nucleic acid sequences. Examples include allele-specific amplification (Kazazian (1989) *J. Lab. Clin. Med.* 114(2):95-96), polymorphism of PCR-amplified fragments (CAPS; Sheffield et al. (1993) *Genomics* 16:325-332), allele-specific ligation (Landegren et al. (1988) *Science* 241:1077-1080), nucleotide extension reactions (Sokolov (1990) *Nucleic Acid Res.* 18:3671), Radiation Hybrid Mapping (Walter et al. (1997) *Nature Genetics* 7:22-28) and Happy Mapping (Dear and Cook (1989) *Nucleic Acid Res.* 17:6795-6807). For these methods, the sequence of a nucleic acid fragment is used to design and produce primer pairs for use in the amplification reaction or in primer extension reactions. The design of such primers is well known to those skilled in the art. In methods employing PCR-based genetic mapping, it may be necessary to identify DNA sequence differences between the parents of the mapping cross in the region corresponding to the instant nucleic acid sequence. This, however, is generally not necessary for mapping methods.

Loss of function mutant phenotypes may be identified for the instant cDNA clones either by targeted gene disruption protocols or by identifying specific mutants for these genes contained in a maize population carrying mutations in all possible genes (Ballinger and Benzer (1989) *Proc. Natl. Acad. Sci USA* 86:9402; Koes et al. (1995) *Proc. Natl. Acad. Sci USA* 92:8149; Bensen et al. (1995) *Plant Cell* 7:75). The latter approach may be accomplished in two ways. First, short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols in conjunction with a mutation tag sequence primer on DNAs prepared from a population of plants in which Mutator transposons or some other mutation-causing DNA element has been introduced (see Bensen, *supra*). The amplification of a specific DNA fragment with these primers indicates the insertion of the mutation tag element in or near the plant gene encoding the instant polypeptides. Alternatively, the instant nucleic acid fragment may be used as a hybridization probe against PCR amplification products generated from the mutation population using the mutation tag sequence primer in conjunction with an arbitrary genomic site primer, such as that for a restriction enzyme site-anchored synthetic adaptor. With either method, a plant containing a mutation in the endogenous gene encoding the instant polypeptides can be identified and obtained. This mutant plant can then be used to determine or confirm the natural function of the instant polypeptides disclosed herein.

#### EXAMPLES

The present invention is further defined in the following Examples, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without

departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

### EXAMPLE 1

#### Composition of cDNA Libraries: Isolation and Sequencing of cDNA Clones

- 5 cDNA libraries representing mRNAs from various barley, corn, rice, soybean and wheat tissues were prepared. The characteristics of the libraries are described below.

**TABLE 2**

cDNA Libraries from Barley, Corn, Rice, Soybean and Wheat

Library	Tissue	Clone
bsh1	Barley ( <i>Hordeum vulgare</i> ) sheath, developing seedling	bsh1.pk0002.f6
cco1n	Corn ( <i>Zea mays</i> ) cob of 67 day old plants grown in green house*	cco1n.pk0005.g3
cdt2c	Corn ( <i>Zea mays</i> ) developing tassel 2	cdt2c.pk002.g1 cdt2c.pk002.116
cr1n	Corn ( <i>Zea mays</i> ) root from 7 day seedlings grown in light*	cr1n.pk0135.e10
csc1c	Corn ( <i>Zea mays</i> ) 20 day seedling (germination under cold stress)	csc1c.pk002.i1
p0031	Corn ( <i>Zea mays</i> ) shoot culture, initiated from seed derived meristems culture was maintained on 273N medium.	p0031.ccmr05rb
p0110	Corn ( <i>Zea mays</i> ) stages V3/V4** leaf tissue minus midrib harvested 4 hours, 24 hours and 7 days after infiltration with salicylic acid, tissues pooled*	p0110.cgsma57r
p0097	Corn ( <i>Zea mays</i> ) stage V9** whorl section (7 cm) from plant infected four times with european corn borer	p0097.cgrad17rc
p0122	Corn ( <i>Zea mays</i> ) pith tissue collected from internode subtending ear node 5 days after pollination	p0122.ckamh70rc
rlr24	Rice ( <i>Oryza sativa</i> ) leaf (15 days after germination) 24 hours after infection of <i>Magaporthe grisea</i> strain 4360-R-62 (AVR2-YAMO); Resistant	rlr24.pk0073.g1
sdp2c	Soybean ( <i>Glycine max</i> ) developing pods 6-7 mm	sdp2c.pk005.o22
ses8w	Soybean ( <i>Glycine max</i> ) mature embryo 8 weeks after subculture	ses8w.pk0028.f3
ssl	Soybean ( <i>Glycine max</i> ) seedling 5-10 day	ssl.pk0036.c10
w11	Wheat ( <i>Triticum aestivum</i> ) leaf 7 day old seedling, light grown	w11.pk0009.c9
w11n	Wheat ( <i>Triticum aestivum</i> ) leaf 7 day old seedling, light grown*	w11n.pk0044.b1
wr1	Wheat ( <i>Triticum aestivum</i> ) root; 7 day old seedling, light grown	wr1.pk0160.d11

Library	Tissue	Clone
wre1n	Wheat ( <i>Triticum aestivum</i> ) root; 7 day old etiolated seedling*	wre1n.pk0043.f9
		wre1n.pk0043.h8
		wre1n.pk0131.g10

\*These libraries were normalized essentially as described in U.S. Patent No. 5,482,845, incorporated herein by reference.

5 \*\*V3, V4 and V9 refer to stages of corn growth. The descriptions can be found in "How a Corn Plant Develops" Special Report No. 48, Iowa State University of Science and Technology Cooperative Extension Service Ames, Iowa, Reprinted February 1993.

cDNA libraries may be prepared by any one of many methods available. For example, the cDNAs may be introduced into plasmid vectors by first preparing the cDNA  
 10 libraries in Uni-ZAP™ XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). The Uni-ZAP™ XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene. Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript. In addition, the cDNAs may be introduced directly into precut Bluescript II SK(+) vectors (Stratagene) using T4 DNA  
 15 ligase (New England Biolabs), followed by transfection into DH10B cells according to the manufacturer's protocol (GIBCO BRL Products). Once the cDNA inserts are in plasmid vectors, plasmid DNAs are prepared from randomly picked bacterial colonies containing recombinant pBluescript plasmids, or the insert cDNA sequences are amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted  
 20 cDNA sequences. Amplified insert DNAs or plasmid DNAs are sequenced in dye-primer sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs"; see Adams et al., (1991) *Science* 252:1651). The resulting ESTs are analyzed using a Perkin Elmer Model 377 fluorescent sequencer.

## EXAMPLE 2

### Identification of cDNA Clones

25 cDNA clones encoding cellulose synthase enzymes were identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol.* 215:403-410; see also [www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)) searches for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS  
 30 translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The cDNA sequences obtained in Example 1 were analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm provided by the National Center for Biotechnology Information (NCBI). The

DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish and States (1993) *Nature Genetics* 3:266-272) provided by the NCBI. For convenience, the P-value (probability) of observing a match of a cDNA sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which represent the negative of the logarithm of the reported P-value. Accordingly, the greater the pLog value, the greater the likelihood that the cDNA sequence and the BLAST "hit" represent homologous proteins.

### EXAMPLE 3

#### Characterization of cDNA Clones Encoding Cellulose Synthase

The BLASTX search using the EST sequences from clones listed in Table 3 revealed similarity of the polypeptides encoded by the cDNAs to cellulose synthase from *Arabidopsis thaliana* (NCBI Identifier No. gi 2827139, gi 2827141, gi 4467125, gi 4886756 and gi 3135611) and *Gossypium hirsutum* (NCBI Identifier No. gi 1706958 and 5081779). Shown in Table 3 are the BLAST results for individual ESTs ("EST"), the sequences of the entire cDNA inserts comprising the indicated cDNA clones ("FIS"), complete gene sequences ("CGS") or contigs assembled from two or more ESTs ("Contig"):

TABLE 3

BLAST Results for Sequences Encoding Polypeptides Homologous to *Arabidopsis thaliana* and *Gossypium hirsutum* Cellulose Synthase

Clone	Status	BLAST pLog Score
bsh1.pk0002.f6	FIS	154.00 (gi 2827139)
Contig composed of: ccoln.pk0005.g3 cdt2c.pk002.g1 cdt2c.pk002.l16 csc1c.pk002.i1 p0031.ccmr05rb p0110.cgsma57r	Contig	>254.00 (gi 2827141)
crln.pk0135.e10	FIS	176.00 (gi 1706958)
p0097.cqrad17rc	CGS	>254.00 (gi 2827141)
p0122.ckamh70rc	CGS	>254.00 (gi 2827141)
rlr24.pk0073.g1	EST	77.70 (gi 4467125)
sdp2c.pk005.o22	FIS	>254.00 (gi 4886756)
ses8w.pk0028.f3	EST	>254.00 (gi 2827139)
ssl.pk0036.c10	EST	>254.00 (gi 2827141)
Contig composed of: wl1.pk0009.c9 wr1.pk0160.d11 wre1n.pk0043.f9	Contig	>254.00 (gi 5081779)

---

wreln.pk0043.h8		
wreln.pk0131.g10		
wlln.pk0044.bl	EST	166.00 (gi 3135611)

---

Figure 1 presents an alignment of the amino acid sequences set forth in SEQ ID NOs:2, 4, 8, 10, 12, 14, 16, 18, 20 and 22 and the *Arabidopsis thaliana* (SEQ ID NOs:23 (gi 2827139), 24 (gi 2827141), 26 (gi 4467125), 27 (gi 4886756) and 29 (gi 3135611) ) and *Gossypium hirsutum* (SEQ ID NOs:25 (gi 1706958) and 28 (gi 5081779)) sequences. The data in Table 4 represents a calculation of the percent identity of the amino acid sequences set forth in SEQ ID NOs:2, 4, 8, 10, 12, 14, 16, 18, 20 and 22 and the *Arabidopsis thaliana* (SEQ ID NOs:23, 24, 26, 27 and 29) and *Gossypium hirsutum* (SEQ ID NOs:25 and 28) sequences.

**TABLE 4**

Percent Identity of Amino Acid Sequences Deduced From the Nucleotide Sequences of cDNA Clones Encoding Polypeptides Homologous to *Arabidopsis thaliana* and *Gossypium hirsutum* Cellulose Synthase

SEQ ID NO.	Percent Identity to
2	82% (gi 2827139)
4	69% (gi 2827141)
6	89% (gi 1706958)
8	70% (gi 2827141)
10	70% (gi 2827141)
12	36% (gi 4467125)
14	86% (gi 4886756)
16	88% (gi 2827139)
18	86% (gi 2827141)
20	87% (gi 5081779)
22	70% (gi 3135611)

Sequence alignments and percent identity calculations were performed using the Megalign program of the LASARGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS*. 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. Sequence alignments and BLAST scores and probabilities indicate that the nucleic acid fragments comprising the instant cDNA clones



encode a substantial portion of a cellulose synthase. These sequences represent the first barley, corn, rice, soybean and wheat sequences encoding cellulose synthase.

#### EXAMPLE 4

##### Expression of Chimeric Genes in Monocot Cells

5 A chimeric gene comprising a cDNA encoding the instant polypeptides in sense orientation with respect to the maize 27 kD zein promoter that is located 5' to the cDNA fragment, and the 10 kD zein 3' end that is located 3' to the cDNA fragment, can be constructed. The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites  
10 (NcoI or SmaI) can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the digested vector pML103 as described below. Amplification is then performed in a standard PCR. The amplified DNA is then digested with restriction enzymes NcoI and SmaI and fractionated on an agarose gel. The appropriate band can be isolated from the gel and combined with a 4.9 kb NcoI-SmaI fragment of the  
15 plasmid pML103. Plasmid pML103 has been deposited under the terms of the Budapest Treaty at ATCC (American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209), and bears accession number ATCC 97366. The DNA segment from pML103 contains a 1.05 kb Sall-NcoI promoter fragment of the maize 27 kD zein gene and a 0.96 kb SmaI-Sall fragment from the 3' end of the maize 10 kD zein gene in the vector  
20 pGem9Zf(+) (Promega). Vector and insert DNA can be ligated at 15°C overnight, essentially as described (Maniatis). The ligated DNA may then be used to transform *E. coli* XL1-Blue (Epicurian Coli XL-1 Blue™; Stratagene). Bacterial transformants can be screened by restriction enzyme digestion of plasmid DNA and limited nucleotide sequence analysis using the dideoxy chain termination method (Sequenase™ DNA Sequencing Kit;  
25 U.S. Biochemical). The resulting plasmid construct would comprise a chimeric gene encoding, in the 5' to 3' direction, the maize 27 kD zein promoter, a cDNA fragment encoding the instant polypeptides, and the 10 kD zein 3' region.

The chimeric gene described above can then be introduced into corn cells by the following procedure. Immature corn embryos can be dissected from developing caryopses  
30 derived from crosses of the inbred corn lines H99 and LH132. The embryos are isolated 10 to 11 days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al. (1975) *Sci. Sin. Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic  
35 proembryoids and embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the primary explant can be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

The plasmid, p35S/Ac (obtained from Dr. Peter Eckes, Hoechst Ag, Frankfurt, Germany) may be used in transformation experiments in order to provide for a selectable marker. This plasmid contains the *Pat* gene (see European Patent Publication 0 242 236) which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The *pat* gene in p35S/Ac is under the control of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*.

The particle bombardment method (Klein et al. (1987) *Nature* 327:70-73) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1  $\mu$ m in diameter) are coated with DNA using the following technique. Ten  $\mu$ g of plasmid DNAs are added to 50  $\mu$ L of a suspension of gold particles (60 mg per mL). Calcium chloride (50  $\mu$ L of a 2.5 M solution) and spermidine free base (20  $\mu$ L of a 1.0 M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After 10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200  $\mu$ L of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30  $\mu$ L of ethanol. An aliquot (5  $\mu$ L) of the DNA-coated gold particles can be placed in the center of a Kapton™ flying disc (Bio-Rad Labs). The particles are then accelerated into the corn tissue with a Biolistic™ PDS-1000/He (Bio-Rad Instruments, Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

For bombardment, the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covered a circular area of about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

Seven days after bombardment the tissue can be transferred to N6 medium that contains glufosinate (2 mg per liter) and lacks casein or proline. The tissue continues to grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to fresh N6 medium containing glufosinate. After 6 weeks, areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the glufosinate-supplemented medium. These calli may continue to grow when sub-cultured on the selective medium.

Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the

tissue can be transferred to regeneration medium (Fromm et al. (1990) *Bio/Technology* 8:833-839).

### EXAMPLE 5

#### Expression of Chimeric Genes in Dicot Cells

5 A seed-specific expression cassette composed of the promoter and transcription terminator from the gene encoding the  $\beta$  subunit of the seed storage protein phaseolin from the bean *Phaseolus vulgaris* (Doyle et al. (1986) *J. Biol. Chem.* 261:9228-9238) can be used for expression of the instant polypeptides in transformed soybean. The phaseolin cassette includes about 500 nucleotides upstream (5') from the translation initiation codon and about  
10 1650 nucleotides downstream (3') from the translation stop codon of phaseolin. Between the 5' and 3' regions are the unique restriction endonuclease sites Nco I (which includes the ATG translation initiation codon), Sma I, Kpn I and Xba I. The entire cassette is flanked by Hind III sites.

The cDNA fragment of this gene may be generated by polymerase chain reaction  
15 (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the expression vector. Amplification is then performed as described above, and the isolated fragment is inserted into a pUC18 vector carrying the seed expression cassette.

20 Soybean embryos may then be transformed with the expression vector comprising sequences encoding the instant polypeptides. To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, can be cultured in the light or dark at 26°C on an appropriate agar medium for 6-10 weeks. Somatic embryos which produce secondary embryos are then excised and  
25 placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos which multiplied as early, globular staged embryos, the suspensions are maintained as described below.

Soybean embryogenic suspension cultures can maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26°C with florescent lights on a 16:8 hour day/night schedule.  
30 Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Klein et al. (1987) *Nature* (London) 327:70, U.S. Patent No. 4,945,050). A DuPont Biolistic™ PDS1000/HE instrument (helium retrofit) can be used  
35 for these transformations.

A selectable marker gene which can be used to facilitate soybean transformation is a chimeric gene composed of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812), the hygromycin phosphotransferase gene from plasmid pJR225

(from *E. coli*; Gritz et al.(1983) *Gene* 25:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The seed expression cassette comprising the phaseolin 5' region, the fragment encoding the instant polypeptides and the phaseolin 3' region can be isolated as a restriction fragment. This fragment can then  
5 be inserted into a unique restriction site of the vector carrying the marker gene.

To 50  $\mu$ L of a 60 mg/mL 1  $\mu$ m gold particle suspension is added (in order): 5  $\mu$ L DNA (1  $\mu$ g/ $\mu$ L), 20  $\mu$ L spermidine (0.1 M), and 50  $\mu$ L  $\text{CaCl}_2$  (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400  $\mu$ L 70%  
10 ethanol and resuspended in 40  $\mu$ L of anhydrous ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five  $\mu$ L of the DNA-coated gold particles are then loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette.  
15 For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into  
25 individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

#### EXAMPLE 6

##### Expression of Chimeric Genes in Microbial Cells

The cDNAs encoding the instant polypeptides can be inserted into the T7 *E. coli* expression vector pBT430. This vector is a derivative of pET-3a (Rosenberg et al. (1987) *Gene* 56:125-135) which employs the bacteriophage T7 RNA polymerase/T7 promoter system. Plasmid pBT430 was constructed by first destroying the EcoR I and Hind III sites in  
35 pET-3a at their original positions. An oligonucleotide adaptor containing EcoR I and Hind III sites was inserted at the BamH I site of pET-3a. This created pET-3aM with additional unique cloning sites for insertion of genes into the expression vector. Then, the Nde I site at the position of translation initiation was converted to an Nco I site using

oligonucleotide-directed mutagenesis. The DNA sequence of pET-3aM in this region, 5'-CATATGG, was converted to 5'-CCCATGG in pBT430.

Plasmid DNA containing a cDNA may be appropriately digested to release a nucleic acid fragment encoding the protein. This fragment may then be purified on a 1% NuSieve GTG™ low melting agarose gel (FMC). Buffer and agarose contain 10 µg/ml ethidium bromide for visualization of the DNA fragment. The fragment can then be purified from the agarose gel by digestion with GELase™ (Epicentre Technologies) according to the manufacturer's instructions, ethanol precipitated, dried and resuspended in 20 µL of water. Appropriate oligonucleotide adapters may be ligated to the fragment using T4 DNA ligase (New England Biolabs, Beverly, MA). The fragment containing the ligated adapters can be purified from the excess adapters using low melting agarose as described above. The vector pBT430 is digested, dephosphorylated with alkaline phosphatase (NEB) and deproteinized with phenol/chloroform as described above. The prepared vector pBT430 and fragment can then be ligated at 16°C for 15 hours followed by transformation into DH5 electrocompetent cells (GIBCO BRL). Transformants can be selected on agar plates containing LB media and 100 µg/mL ampicillin. Transformants containing the gene encoding the instant polypeptides are then screened for the correct orientation with respect to the T7 promoter by restriction enzyme analysis.

For high level expression, a plasmid clone with the cDNA insert in the correct orientation relative to the T7 promoter can be transformed into *E. coli* strain BL21(DE3) (Studier et al. (1986) *J. Mol. Biol.* 189:113-130). Cultures are grown in LB medium containing ampicillin (100 mg/L) at 25°C. At an optical density at 600 nm of approximately 1, IPTG (isopropylthio-β-galactoside, the inducer) can be added to a final concentration of 0.4 mM and incubation can be continued for 3 h at 25°. Cells are then harvested by centrifugation and re-suspended in 50 µL of 50 mM Tris-HCl at pH 8.0 containing 0.1 mM DTT and 0.2 mM phenyl methylsulfonyl fluoride. A small amount of 1 mm glass beads can be added and the mixture sonicated 3 times for about 5 seconds each time with a microprobe sonicator. The mixture is centrifuged and the protein concentration of the supernatant determined. One µg of protein from the soluble fraction of the culture can be separated by SDS-polyacrylamide gel electrophoresis. Gels can be observed for protein bands migrating at the expected molecular weight.

#### EXAMPLE 7

##### Evaluating Compounds for Their Ability to Inhibit the Activity of Cellulose Synthase

The polypeptides described herein may be produced using any number of methods known to those skilled in the art. Such methods include, but are not limited to, expression in bacteria as described in Example 6, or expression in eukaryotic cell culture, *in planta*, and using viral expression systems in suitably infected organisms or cell lines. The instant

polypeptides may be expressed either as mature forms of the proteins as observed *in vivo* or as fusion proteins by covalent attachment to a variety of enzymes, proteins or affinity tags. Common fusion protein partners include glutathione S-transferase ("GST"), thioredoxin ("Trx"), maltose binding protein, and C- and/or N-terminal hexahistidine polypeptide ("His)<sub>6</sub>"). The fusion proteins may be engineered with a protease recognition site at the fusion point so that fusion partners can be separated by protease digestion to yield intact mature enzyme. Examples of such proteases include thrombin, enterokinase and factor Xa. However, any protease can be used which specifically cleaves the peptide connecting the fusion protein and the enzyme.

Purification of the instant polypeptides, if desired, may utilize any number of separation technologies familiar to those skilled in the art of protein purification. Examples of such methods include, but are not limited to, homogenization, filtration, centrifugation, heat denaturation, ammonium sulfate precipitation, desalting, pH precipitation, ion exchange chromatography, hydrophobic interaction chromatography and affinity chromatography, wherein the affinity ligand represents a substrate, substrate analog or inhibitor. When the instant polypeptides are expressed as fusion proteins, the purification protocol may include the use of an affinity resin which is specific for the fusion protein tag attached to the expressed enzyme or an affinity resin containing ligands which are specific for the enzyme. For example, the instant polypeptides may be expressed as a fusion protein coupled to the C-terminus of thioredoxin. In addition, a (His)<sub>6</sub> peptide may be engineered into the N-terminus of the fused thioredoxin moiety to afford additional opportunities for affinity purification. Other suitable affinity resins could be synthesized by linking the appropriate ligands to any suitable resin such as Sepharose-4B. In an alternate embodiment, a thioredoxin fusion protein may be eluted using dithiothreitol; however, elution may be accomplished using other reagents which interact to displace the thioredoxin from the resin. These reagents include  $\beta$ -mercaptoethanol or other reduced thiol. The eluted fusion protein may be subjected to further purification by traditional means as stated above, if desired. Proteolytic cleavage of the thioredoxin fusion protein and the enzyme may be accomplished after the fusion protein is purified or while the protein is still bound to the ThioBond™ affinity resin or other resin.

Crude, partially purified or purified enzyme, either alone or as a fusion protein, may be utilized in assays for the evaluation of compounds for their ability to inhibit enzymatic activation of the instant polypeptides disclosed herein. Assays may be conducted under well known experimental conditions which permit optimal enzymatic activity. For example, assays for cellulose synthase activity are presented in WO 98/18949 and WO 98/00549.

### CLAIMS

What is claimed is:

1. An isolated nucleic acid fragment comprising at least 900 nucleotides, wherein the nucleic acid fragment encodes a cellulose synthase comprising a member selected from the group consisting of:
  - (a) an isolated nucleic acid fragment encoding an amino acid sequence that is at least 90% identical to the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, 6, 12, 14, 16, 18, 20 and 22;
  - (b) an isolated nucleic acid fragment that is complementary to (a).
2. The isolated nucleic acid fragment of Claim 1 wherein nucleic acid fragment is a functional RNA.
3. The isolated nucleic acid fragment of Claim 1 wherein the nucleotide sequence of the fragment comprises the sequence set forth in a member selected from the group consisting of SEQ ID NO:1, 5, 11, 13, 15, 17, 19 and 21.
4. A chimeric gene comprising the nucleic acid fragment of Claim 1 operably linked to suitable regulatory sequences.
5. A transformed host cell comprising the chimeric gene of Claim 4.
6. A cellulose synthase polypeptide comprising all or a substantial portion of the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, 6, 12, 14, 16, 18, 20 and 22.
7. An isolated nucleic acid fragment encoding a cellulose synthase comprising a member selected from the group consisting of:
  - (a) an isolated nucleic acid fragment encoding an amino acid sequence that is functionally active polypeptide and at least 80% identical to the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:4, 8 and 10 ;
  - (b) an isolated nucleic acid fragment that is complementary to (a).
8. The isolated nucleic acid fragment of Claim 7 wherein nucleic acid fragment is a functional RNA.
9. The isolated nucleic acid fragment of Claim 7 wherein the nucleotide sequence of the fragment comprises the sequence set forth in a member selected from the group consisting of SEQ ID NO:3, 7 and 9.
10. A chimeric gene comprising the nucleic acid fragment of Claim 7 operably linked to suitable regulatory sequences.
11. A transformed host cell comprising the chimeric gene of Claim 10.

12. A cellulose synthase polypeptide comprising all or a substantial portion of the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:4, 8, 10.

5 13. A method of altering the level of expression of a cellulose synthase in a host cell comprising:

(a) transforming a host cell with the chimeric gene of any of Claims 4 and 10; and

(b) growing the transformed host cell produced in step (a) under conditions that are suitable for expression of the chimeric gene

10 wherein expression of the chimeric gene results in production of altered levels of a cellulose synthase in the transformed host cell.

14. A method of obtaining a nucleic acid fragment encoding all or a substantial portion of the amino acid sequence encoding a cellulose synthase comprising:

15 (a) probing a cDNA or genomic library with the nucleic acid fragment of any of Claims 1 and 7;

(b) identifying a DNA clone that hybridizes with the nucleic acid fragment any of of Claims 1 and 7;

(c) isolating the DNA clone identified in step (b); and

20 (d) sequencing the cDNA or genomic fragment that comprises the clone isolated in step (c)

wherein the sequenced nucleic acid fragment encodes all or a substantial portion of the amino acid sequence encoding a cellulose synthase.

15. A method of obtaining a nucleic acid fragment encoding a substantial portion of an amino acid sequence encoding a cellulose synthase comprising:

25 (a) synthesizing an oligonucleotide primer corresponding to a portion of the sequence set forth in any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19 and 21; and

30 (b) amplifying a cDNA insert present in a cloning vector using the oligonucleotide primer of step (a) and a primer representing sequences of the cloning vector

wherein the amplified nucleic acid fragment encodes a substantial portion of an amino acid sequence encoding a cellulose synthase.

16. The product of the method of Claim 14.

17. The product of the method of Claim 15.

35 18. A method for evaluating at least one compound for its ability to inhibit the activity of a cellulose synthase, the method comprising the steps of:



- 5
- 10
- (a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding a cellulose synthase, operably linked to suitable regulatory sequences;
  - (b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of the cellulose synthase encoded by the operably linked nucleic acid fragment in the transformed host cell;
  - (c) optionally purifying the cellulose synthase expressed by the transformed host cell;
  - (d) treating the cellulose synthase with a compound to be tested; and
  - (e) comparing the activity of the cellulose synthase that has been treated with a test compound to the activity of an untreated cellulose synthase, thereby selecting compounds with potential for inhibitory activity.

Figure 1

```

1
SEQ ID NO: 2 -----
SEQ ID NO: 4 RAAQAQRNKGKQPPEEQKLASVSLP--LPHSRFIPFPPRRRYRRR---RTHACPG----I
SEQ ID NO: 6 -----
SEQ ID NO: 8 HSSYTKSRSSLAQPRAAPRQAQPPP--ATAACACERSPRPGDQRRGGLRAFRCAAAAGFV
SEQ ID NO: 10 -----
SEQ ID NO: 12 RCS---RRWTCSSPPPTPTRSRSPRRTP-----
SEQ ID NO: 14 -----
SEQ ID NO: 16 -----
SEQ ID NO: 18 -----
SEQ ID NO: 20 -----
SEQ ID NO: 22 -----
SEQ ID NO: 23 -----
SEQ ID NO: 24 MNTGGR-----
SEQ ID NO: 25 -----
SEQ ID NO: 26 MASTPPQTSKKVRNNSGSGQTVKFARRTSSGRYVSLR-RDNIELSGELSGDYSNYTVHIP
SEQ ID NO: 27 -----
SEQ ID NO: 28 -----
SEQ ID NO: 29 R---PR-----

61
SEQ ID NO: 2 -----
SEQ ID NO: 4 W-RSGSARG----ME-ASAGLVAGSHNRNELV-VIRRDGEPGPKP--MDQRNGQVCQI--
SEQ ID NO: 6 -----
SEQ ID NO: 8 RERDPAGRGGGPEME-ASAGLVAGSHNRNELV-VIRRDRESGAAGGGAARRAEAPCQI--
SEQ ID NO: 10 -----
SEQ ID NO: 12 -----ME-ASAGLVAGSHNRNELV-VIRRDGDPGPKP--PREQNGQVCQI--
SEQ ID NO: 14 -----C-----
SEQ ID NO: 16 -----ME-ASAGLVAGSHNRNELV-VIHGHEEP--KA--LKNLDGQVCEI--
SEQ ID NO: 18 -----
SEQ ID NO: 20 -----
SEQ ID NO: 22 -----
SEQ ID NO: 23 -----ME-ASAGLVAGSYRRNELV-RIRHESDGGTKP--LKNMNGQICQI--
SEQ ID NO: 24 -----LIAGSHNRNEFV-LI--NADESARIRSVQELSGQTCQI--
SEQ ID NO: 25 -----
SEQ ID NO: 26 PTPDNQPMATKAEEQYVNSLFTGGFNSVTRAHMDKVIDSDVTHPQMAAGKSSCAMP
SEQ ID NO: 27 -----ME-ASAGLVAGSHNRNELV-VIHNEEP--KP--LKNLDGQFCEI--
SEQ ID NO: 28 -----
SEQ ID NO: 29 -----LIAGSHNRNEFV-LI--NADENARIRSVQELSGQTCQI--

121
SEQ ID NO: 2 -----
SEQ ID NO: 4 CGDDVGRNPDGEPFVACNECAFPICRDCYERREGTQNCQPCKTRYKRLKGCARVPGD-
SEQ ID NO: 6 -----
SEQ ID NO: 8 CGDEVGVGFDGEPFVACNECAFPVCRACYERREGSQACPQCRTRYKRLKGCPRVAGD-
SEQ ID NO: 10 CGDDVGLAPGGDPFVACNECAFPVCRDCYERREGTQNCQPCKTRYKRLKGCQRTVTD-
SEQ ID NO: 12 -----CPY-----
SEQ ID NO: 14 CGDGVGLTVGDGLFVACNECGFPVCRPCYERREGSHLCPQCKTRYKRLKGSPRVEGDD
SEQ ID NO: 16 -----
SEQ ID NO: 18 -----
SEQ ID NO: 20 -----
SEQ ID NO: 22 -----
SEQ ID NO: 23 CGDDVGLAETGDVVFACNECAFPVCRPCYERKDGTOCCPQCKTRFRHRHGS PRVEGDE
SEQ ID NO: 24 CGDEIELTVSSELFVACNECAFPVCRPCYERREGNQACFPQCKTRYKRIKGS PRVDGDD
SEQ ID NO: 25 -----
SEQ ID NO: 26 CDGNVMKDERGKDVMP-ECRFKICRDCFMDAQE-TGLCPGCKEQYK-----

```

Figure 1 (cont'd.)

```

SEQ ID NO:27 CGDQIGLTVEGDLFVACNECGFPACRPCYEYERREGTQNCPCQCKTRYKRLRGS PRVEGDE
SEQ ID NO:28 -----
SEQ ID NO:29 CRDEIELTVDGEPFVACNECAFPVCRPCYEYERREGNQACPQCKTRFKRLKGS PRVEGD-

181
SEQ ID NO:2 ----- 240
SEQ ID NO:4 EEEDGVDDLENEFNWSDK----HDSQYLAESMLHAHMSYG-RGADLDGVPQPFHPINVP
SEQ ID NO:6 -----
SEQ ID NO:8 EEEDGVDDLEGEFGLQDGAHEDDPQYVAESMLRAQMSYG-RGGDAH---PGFSPVPNVP
SEQ ID NO:10 EEEDGVDDLNEFNW-DG----HDSQVAESMLYGHMSYG-RGGDPNGAPQAFQLNPNVP
SEQ ID NO:12 -----
SEQ ID NO:14 DEEDV-DDIEHEFNIDEQKNKHGQ---VAEAMLHGRMSYG--RGPEDDDNSQFPTPIAG
SEQ ID NO:16 -----
SEQ ID NO:18 -----
SEQ ID NO:20 -----
SEQ ID NO:22 -----
SEQ ID NO:23 DEDDV-DDIENEFNYAQGANKARH---QRHGE---EFSSS--SRHESQPIPLLTHGHTVS
SEQ ID NO:24 EEEEDIDDLLEYFD-----HGMDPEHAAEAALSSRLNTG--RGLDLSAPPG-----SQIP
SEQ ID NO:25 -----
SEQ ID NO:26 -----IGDLDD-----TPDYSSGALPLPAPG-----
SEQ ID NO:27 DEEDI-DDIEYEFNIEHEQDKKH---SAEAMLYGKMSYG--RGPEDDENGFRP-PVIAG
SEQ ID NO:28 -----
SEQ ID NO:29 EEEDDIDDLNEFEYGN---NGIGFDQVSEGMSISRRNSGFPQSDLSAPPG-----SQIP

241
SEQ ID NO:2 ----- 300
SEQ ID NO:4 LLTNGQMVDIPPDQHALVPSFV---GGGGKRIHPLPYADPNLPVQPRSM DPSKDLAAYG
SEQ ID NO:6 -----
SEQ ID NO:8 LLTNGQMVDIPPEQHALVPSYMSGGGGGGKRIHPLPFADPNLPVQPRSM DPSKDLAAYG
SEQ ID NO:10 LLTNGQMVDIPPEQHALVPSFM---GGGGKRIHPLPYADPSLPVQPRSM DPSKDLAAYG
SEQ ID NO:12 -----
SEQ ID NO:14 GRSR-----PVSGEFPISSNAYGDQMLSSSLHKRVHPYPVSEPGSARW-----DEKXKDG
SEQ ID NO:16 -----
SEQ ID NO:18 -----
SEQ ID NO:20 -----
SEQ ID NO:22 LLTNGQMVDIPPEQHALVPSYMSGGGGGGKRIHPLPFADPNLPVQPRSM DPSKDLAAYG
SEQ ID NO:23 GEIRTPDTQSVRTTSGPLGPSDRNAISSPYIDPR-QPVPVRIVDPSK-----DLNSYG
SEQ ID NO:24 LLTYCDEDADMYSRHALIVP--PS-TGYGNRVYPAPFTDSSAPPQARSMPVQKDIAEYG
SEQ ID NO:25 -----
SEQ ID NO:26 -----KDQRGNNNNMSMMKRNQNGEFDHNRWLF-----ETQGTYG
SEQ ID NO:27 GHS-----GEFPVGG-GYGNG--EHGLHKRVHPYPSSEAGS-----EGG
SEQ ID NO:28 -----
SEQ ID NO:29 LLTYGDEDVEISSDRHALIVP--PSLGGHGNRVHPVSLSDPTVAHRRLMVPQKDLAVYG

301
SEQ ID NO:2 ----- 360
SEQ ID NO:4 YGSVAWKERMESWKQKQ-ERMHQTNRNDGGGD-----DGDDADLPLM-DEARQPLSR
SEQ ID NO:6 -----
SEQ ID NO:8 YGSVAWKERMESWKQKQ-ERLQHVRESEGGGDW-----DGDDADLPLM-DEARQPLSR
SEQ ID NO:10 YGSVAWKERMENWKQRQ-ERMHQTNRNDGGGD-----DGDDADLPLM-DEARQPLSR
SEQ ID NO:12 -----
SEQ ID NO:14 -----WKDRMDDWKLOQG-----NLGPEPDEDPAAML-DEARQPLSR
SEQ ID NO:16 -----
SEQ ID NO:18 -----HE-----
SEQ ID NO:20 -----
SEQ ID NO:22 YGSVAWKERMESWKQKQ-ERLQHVRESEGGGDW-----DGDDADLPLM-DEARQPLSR
SEQ ID NO:23 LGNVDWKERVEGWKLQEKNLQMTGKYHEGKGG-EIEGTGSNGEELQM-ADDTRLPMR

```

Figure 1 (cont'd.)

SEQ ID NO:24 YGSVAWKDRMEVWKRRQGEKLVKHEGGNNGRGSN--DDELDLDDPDMPM--DEGRQPLSR  
 SEQ ID NO:25 -----  
 SEQ ID NO:26 YGNAYWP-----QDEMYGD-----DMDEGMRGGMVETADKPWRPLSR  
 SEQ ID NO:27 -----WRERMDDWKLOHG-----NLGPEPDDDPDEMGLI--DEARQPLSR  
 SEQ ID NO:28 -----  
 SEQ ID NO:29 YGSVAWKDRMEEWKRKQNEKLQVVRHEGDP-----DFEDGDDADFFMM--DEGRQPLSM

361

420

SEQ ID NO:2 -----  
 SEQ ID NO:4 KIPLPSSQINPYRMIIIRLVVLCFFFHYRVMHPVDAFALWLISVICEIWFAMSWILDQ  
 SEQ ID NO:6 -----  
 SEQ ID NO:8 KVPISSSRINPYRMIIIRLVVLCFFFHYRVMHPAKDAFALWLISVICEIWFAMSWILDQ  
 SEQ ID NO:10 KIPLPSSQINPYRMIIIRLVVLCFFFHYRVMHPVDAFALWLISVICEIWFAMSWILDQ  
 SEQ ID NO:12 -----  
 SEQ ID NO:14 KVPIASSKINPYRMVIVARLVILAFFLRYRLMNPVHDALGLWLTSIICEIWFAFSWILDQ  
 SEQ ID NO:16 -----  
 SEQ ID NO:18 -----LHPVNDAYGLWLTSVICEIWFVSWIMDQ  
 SEQ ID NO:20 -----  
 SEQ ID NO:22 KVPISSSRINPYRMIIIRLVVLCFFFHYRVMHPAKDAFALWLISVICEIWFAMSWILDQ  
 SEQ ID NO:23 VVIPSSRLTPYRVVILRLIILCFFLQYRTHPVKNAYPLWLTSVICEIWFVAFSWILDQ  
 SEQ ID NO:24 KLPISSSRINPYRMILLCRLAILGLFFHYRILHPVNDAYGLWLTSVICEIWFVAFSWILDQ  
 SEQ ID NO:25 -----  
 SEQ ID NO:26 RIPIAAIISPYRLLIVIRFVLCFFLTWRIRNPNEDAIWLWMSIICELWFGFSWILDQ  
 SEQ ID NO:27 KVPIASSKINPYRMVIVARLVILAVFLRYRLMNPVHDALGLWLTSVICEIWFVAFSWILDQ  
 SEQ ID NO:28 -----  
 SEQ ID NO:29 KIPIKSSKINPYRMLIVLRLVILGLFFHYRILHPVKDAYALWLISVICEIWFVAFSWILDQ

421

480

SEQ ID NO:2 -----  
 SEQ ID NO:4 FPKWFPIERETYLDRSLRFDKEGHPS-----QLAPVDFFVSTVDPLKEPPLVTANTVLS  
 SEQ ID NO:6 -----  
 SEQ ID NO:8 FPKWLPPIERETYLDRSLRFDKEGQPS-----QLAPIDFFVSTVDPTKEPPLVTANTVLS  
 SEQ ID NO:10 FPKWFPIERETYLDRSLRFDKEGQPS-----QLAPIDFFVSTVDPLKEPPLVTANTVLS  
 SEQ ID NO:12 -----  
 SEQ ID NO:14 FPKWFPIDRETYLDRSLIRYEREGEPN-----MLAPVDVFVSTVDPMKEPPLVTANTVLS  
 SEQ ID NO:16 -----  
 SEQ ID NO:18 FPKWYPIQRETYLDRSLRYEKEGKPS-----ELSSVDVFVSTVDPMKEPPLITANTVLS  
 SEQ ID NO:20 -----  
 SEQ ID NO:22 FPKWFPIERETYLDRSLRFDKEGQPS-----QLAPIDFFVSTVDPTKEPPLVTANTVLS  
 SEQ ID NO:23 FPKWYPINRETYLDRSLAIRYDRDGEPS-----QLVPVDVFVSTVDPLKEPPLVTANTVLS  
 SEQ ID NO:24 FPKWYPIERETYLDRSLRYEKEGKPS-----GLAPVDVFVSTVDPLKEPPLITANTVLS  
 SEQ ID NO:25 -----  
 SEQ ID NO:26 IPKLCPINRSTDLEVLDRDKFDMSPSNPTGRSDLPGLDFVSTADPEKEPPLVTANTVLS  
 SEQ ID NO:27 FPKWFPIERETYLDRSLRYEREGEPN-----MLAPVDVFVSTVDPLKEPPLVTANTVLS  
 SEQ ID NO:28 -----  
 SEQ ID NO:29 FPKWYPIERETYLDRSLRYEKEGKPS-----GLSPVDVFVSTVDPLKEPPLITANTVLS

481

540

SEQ ID NO:2 -----  
 SEQ ID NO:4 ILSVDYPVDKVSCYVSDGAAMLTFEALSETSEFAKKWVPFCKRYSLEPRAPEWYFQ--  
 SEQ ID NO:6 -----H-----  
 SEQ ID NO:8 ILSVDYPVEKVSCYVSDGAAMLTFEALSETSEFAKKWVPFCKRYSLEPRAPEWYFQ--  
 SEQ ID NO:10 ILSVDYPVDKVSCYVSDGAAMLTFEALSETSEFAKKWVPFCKRYSLEPRAPEWYFQ--  
 SEQ ID NO:12 ILAAGYPAGKVTCTYISDDAGAETRNAVVEAARFAALWVSFCRKHGVEPRNLEAYFNAGE  
 SEQ ID NO:14 ILAMDYPVDKISCYISDDGASMCTFESLSETAEFARKWVPFCKKFSIEPRAPEMYFSE--  
 SEQ ID NO:16 -----  
 SEQ ID NO:18 ILAVDYPVDKVACYVSDGAAMLTFEALSETSEFARRWVPFCKKYNIEPRAPEWYFQ--

Figure 1 (cont'd.)

```

SEQ ID NO:20 -----
SEQ ID NO:22 ILSVDYPVEKVSCYVSDGGAAMLTFEALSETSEFAKKWVPFSKKFNIEPRAPEWYFQQ--
SEQ ID NO:23 ILSVDYPVDKVACYVSDDGSAAMLTFEALSETAEFAKKWVPFCKKFNIEPRAPEFYFAQ--
SEQ ID NO:24 ILAVDYPVDKVACYVSDGGAAMLTFEALSDTAEFARKWVPFCKKFNIEPRAPEWYFSQ--
SEQ ID NO:25 -----RRWVPFCKKHNVFPRAPEFYFNE--
SEQ ID NO:26 ILAVDYPVEKVSCYLSDDGGALLSFEAMAEASFADLWVPFCRKHNIERNPDSYFSL--
SEQ ID NO:27 ILAMDYPVEKISCYVSDDGASMLTFEALSETAEFAKKWVPFCKKFSIEPRAPEMYFTL--
SEQ ID NO:28 ----DYPVEKVSCYVSDGGAAMLTFEALSETSEFAKKWVPFCKKYNIEPRAPEWYFAQ--
SEQ ID NO:29 ILAVDYPVDKVACYVSDGGAAMLTFEALSETAEFAKKWVPFCKKYCIEPRAPEWYFCH--

541
SEQ ID NO:2 -----600
SEQ ID NO:4 -----KIDYLDKDKVAPNFRERRAMKREYEEFKVRINALVAKAQ-----
SEQ ID NO:6 -----
SEQ ID NO:8 -----KIDYLDKDKVAASFVRERRAMKREYEEFKVRINALVAKAQ-----
SEQ ID NO:10 -----KIDYLDKDKVAANFVRERRAMKREYEEFKVRINALVAKAQ-----
SEQ ID NO:12 GGGGKAKVVARGSY-RGMAWPELVDRRRVRREYEEMLRIDALQAADARRR-----
SEQ ID NO:14 -----KIDYLDKDKVQPTFVKERRAMKREYEEFKVRINALVAKAQ-----
SEQ ID NO:16 -----AKAQ-----
SEQ ID NO:18 -----KMDYLKNKVHAPAFVRERRAMKRDYEEFKVRINSLVATAQ-----
SEQ ID NO:20 -----
SEQ ID NO:22 -----KIDYLDKDKVAASFVRERRAMKREYEEFKVRINALVAKAQ-----
SEQ ID NO:23 -----KIDYLDKDKIQPSFVKERRAMKREYEEFKVRINALVAKAQ-----
SEQ ID NO:24 -----KMDYLKNKVHAPAFVRERRAMKRDYEEFKVKINALVATAQ-----
SEQ ID NO:25 -----KIDYLDKDKVHPSFVKERRAMKREYEEFKVRINALVAKAQ-----
SEQ ID NO:26 -----KIDPTKNKSRIDFVKDRRKIKREYDEFKVRINGLPDSIRRRSDAFNAREE
SEQ ID NO:27 -----KVDYLDQDKVHPTFVKERRAMKREYEEFKVRINAQVAKAS-----
SEQ ID NO:28 -----KIDYLDKDKVQTSFVKERRAMKREYEEFKVRVNLVAKAQ-----
SEQ ID NO:29 -----KMDYLKNKVHAPAFVRERRAMKRDYEEFKVKINALVATAQ-----

601
SEQ ID NO:2 -----660
SEQ ID NO:4 -----KVPEEGWTMQDGTWPWG-----NNVRDHPGMIQVFL---
SEQ ID NO:6 -----
SEQ ID NO:8 -----KVPEEGWTMQDGSPWPG-----NNVRDHPGMIQVFL---
SEQ ID NO:10 -----KVPEEGWTMQDGTWPWG-----NNVRDHPGMIQVFL---
SEQ ID NO:12 -----RRGAADDHAGVVQVLIDFA
SEQ ID NO:14 -----KVPOGGWIMQDGTWPWG-----NNTKDHPGMIQVFL---
SEQ ID NO:16 -----KMPEEGWTMQDGTWPWG-----NNPRDHPGMIQVFL---
SEQ ID NO:18 -----KVPEDGWTMQDGTWPWG-----NNVRDHPGMIQVFL---
SEQ ID NO:20 -----
SEQ ID NO:22 -----KVPEEGWTMQDGSPWPG-----
SEQ ID NO:23 -----KIPEEGWTMQDGTWPWG-----NNTRDHPGMIQVFL---
SEQ ID NO:24 -----KVPEEGWTMQDGTWPWG-----NNVRDHPGMIQVFL---
SEQ ID NO:25 -----KKPEEGWVMQDGTWPWG-----NNTRDHPGMIQVYL---
SEQ ID NO:26 MKALKQMRESGGDPTFPVKPKATW-MADGTHWPGTWAASTREHSKGDHAGILQVMLKPP
SEQ ID NO:27 -----KVPLEGWIMQDGTWPWG-----NNTKDHPGMIQVFL---
SEQ ID NO:28 -----KVPEEGWIMQDGTWPWG-----NNTRDHPGMIQVFL---
SEQ ID NO:29 -----KVPEDGWTMQDGTWPWG-----NSVRDHPGMIQVFL---

661
SEQ ID NO:2 -----720
SEQ ID NO:4 -----G-QSGGHDVE----GNELPRLVYVSREKRPGYNHKKAGAMNALVRVSAVLTNA
SEQ ID NO:6 -----
SEQ ID NO:8 -----G-QSGGRDVE----GNELPRLVYVSREKRPGYNHKKAGAMNALVRVSAVLSNA
SEQ ID NO:10 -----G-QSGGLDCE----GNELPRLVYVSREKRPGYNHKKAGAMNALVRVSAVLTNA
SEQ ID NO:12 GSVPQLGVANGSKLIDVASVDVCLPALVYVCREKRRGHAHHRKAGAMNA-----

```

Figure 1 (cont'd.)

```

SEQ ID NO:14 -----G-SSGGLDTE----GNQLPRLVYVSREKRPGFQHHKKAGAMNALVRVSAVLTNA
SEQ ID NO:16 -----G-HSGGLDTD----GNELPRLVYVSREKRPGFQHHKKAGAMNALIRVSAVLTNG
SEQ ID NO:18 -----G-QDGV RDVE----GNELPRLVYVSREKRPGFQHHKKAGAMNALVRASAITNA
SEQ ID NO:20 -----
SEQ ID NO:22 -----
SEQ ID NO:23 -----G-HSGGLDTD----GNELPRLIYVSREKRPGFQHHKKAGAMNALIRVSAVLTNG
SEQ ID NO:24 -----G-HSGV RDTD----GNELPRLVYVSREKRPGFQHHKKAGAMNSLIRVSAVLSNA
SEQ ID NO:25 -----G-SAGALDVD----GKELPRLVYVSREKRPQYQHHKKAGAENALVRVSAVLTNA
SEQ ID NO:26 SSDPLIG-NSDDKVIDFSDTDTRLPMFVYVSREKRPQYDHNKKAGAMNALVRASAILSNG
SEQ ID NO:27 -----G-HSGGFDVE----GHELPLRVYVSREKRPGFQHHKKAGAMNALVRVAGVLTNA
SEQ ID NO:28 -----G-QSGGLDAE----GNELPRLVYVSREKRPGFQHHKKAGAMNALVRVSAVLTNG
SEQ ID NO:29 -----G-SDGV RDVE----NNELPRLVYVSREKRPGFQHHKKAGAMNSLIRVSGVLSNA

```

721

780

```

SEQ ID NO:2 -----
SEQ ID NO:4 PYLLNLDCDHYINNSKAIKEAMCFMMDPLLGGK-----VCYVQFPQRFDGIDRHDRYAN
SEQ ID NO:6 -----
SEQ ID NO:8 AYLLNLDCDHYINNSKAIKEAMCFMMDPLVGKK-----VCYVQFPQRFDGIDKNDRYAN
SEQ ID NO:10 PYLLNLDCDHYINNSKAIKEAMCFMMDPLLGGK-----VCYVQFPQRFDGIDRHDRYAN
SEQ ID NO:12 PFILDLD CDYVNNSQLRAGICFMIERG GGAEDAGAVAFVQFPQRV DGVDPGDRYAN
SEQ ID NO:14 PFMLNLDCDHYVNNSKAAREAMCFLMDPQTGKK-----VCYVQFPQRFDGIDTHDRYAN
SEQ ID NO:16 AYLLNVDCDHYFNNSKALKEAMCFMMDPVLGGK-----TCYVQFPQRFDGIDLHDRYAN
SEQ ID NO:18 PYLLNVDCDHYINNSKALREAMCFMMDPQLGGK-----VCYVQFPQRFDGIDRHDRYSN
SEQ ID NO:20 -----EAMCFLMDPNLGPQ-----VCYVQFPQRFDGIDRNDRYAN
SEQ ID NO:22 -----
SEQ ID NO:23 AYLLNVDCDHYFNNSKAIKEAMCFMMDPAIGKK-----CCYVQFPQRFDGIDLHDRYAN
SEQ ID NO:24 PYLLNVDCDHYINNSKAIRESMCFMMDPQSGKK-----VCYVQFPQRFDGIDRHDRYSN
SEQ ID NO:25 PFILNLDCDHYINNSKAMREAMCFLMDPQFGKK-----LCYVQFPQRFDGIDRHDRYAN
SEQ ID NO:26 PFILNLDCDHYIYNCKAVREGMCFMMDRG-GED-----ICYIQFPQRFEGIDPSDRYAN
SEQ ID NO:27 PFMLNLDCDHYVNNSKAVREAMCFLMDPQIGKK-----VCYVQFPQRFDGIDTNDRYAN
SEQ ID NO:28 AFLNLDCDHYINNSKALREAMCFLMDPNLGGK-----VCYVQFPQRFDGIDRNDRYAN
SEQ ID NO:29 PYLLNVDCDHYINNSKALREAMCFMMDPQSGKK-----ICYVQFPQRFDGIDRHDRYSN

```

781

840

```

SEQ ID NO:2 -----
SEQ ID NO:4 RNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAP---KTKKPPSRTCNCWPKWCI
SEQ ID NO:6 -----
SEQ ID NO:8 RNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAP---KTKKPPSRTCNCWPKWCL
SEQ ID NO:10 RNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGYDAP---KTKKPPSRTCNCWPKWCF
SEQ ID NO:12 HNRVLFDC TELGLDGLQGPIYVGTGCLFRRVALYSVDLPR-----
SEQ ID NO:14 RNTVFFDINMKGLDGIQGPVYVGTGCVFRRQALYGYNPPKGP--RPMVSCDC-----
SEQ ID NO:16 RNIVFFDINMKGQDGVQGPVYVGTGCCFNQALYGYDPVLTEEDLE-----PNIIV
SEQ ID NO:18 RNVVFFDINMKGLDGIQGPIYVGTGCVFRRYALYGYDAP---AKKKPPSKTCNCWPKWCC
SEQ ID NO:20 RNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPKAK--K-----PGFLA
SEQ ID NO:22 -----
SEQ ID NO:23 RNIVFFDINMKGLDGIQGPVYVGTGCCFNQALYGYDPVLTEEDLE-----PNIIV
SEQ ID NO:24 RNVVFFDINMKGLDGIQGPIYVGTGCVFRRQALYGFDAF---KKKKPPGKTCNCWPKWCC
SEQ ID NO:25 RNVVFFDINMLGLDGLQGPIYVGTGCVFNQALYGYDPPVSEKRPK---MTCDCWPSWCC
SEQ ID NO:26 NNTVFFDGNMRALDGVQGPVYVGTGTMFRRFALYGFDP-----
SEQ ID NO:27 RNTVFFDINMKGLDGIQGPVYVGTGCVFKRQALYGYEPPKGP--RPMISCGC-----
SEQ ID NO:28 RNTVFFDINLRGLDGIQGPVYVGTGCVFNRTALYGYEPLKPKHRK-----TGILS
SEQ ID NO:29 RNVVFFDINMKGLDGLQGPIYVGTGCVFRRQALYGFDAF---KKKKGPRKTCNCWPKWCL

```

Figure 1 (cont'd.)

	841	900
SEQ ID NO:2	-----	-----
SEQ ID NO:4	CCCCFGNRKTKKKTKTSKP-----	KFEKIKKLF-KKKENQAPAYALGEIDEA--APG-
SEQ ID NO:6	-----	-----
SEQ ID NO:8	SCCC--SRNKNKKKTTKP-----	KTEKKRRLFFKKAENPSPAYALGEIDEG--APG-
SEQ ID NO:10	CCCCFGNRKQKK---TTKP-----	KTEKKKLLFFKKEENQSPAYALGEIDEA--APG-
SEQ ID NO:12	-----	-----
SEQ ID NO:14	-CPCFGSRKKYKE-----	-----KNDANGEAASLKG
SEQ ID NO:16	-KSCCGSRKKGKGGNKK-----	YS-DKKKAMGR--TESTVPIFNMEDIEEGVEG--Y
SEQ ID NO:18	LCC--GSRKKKN---ANS-----	KKEKKRKV--KHSEASKQIHALENIEAGN--EG-
SEQ ID NO:20	-SLCXG-KKKASKSKKR-----	SSDKKKS NKH--VDSSVPVFNLEDIEEGVEGAGF
SEQ ID NO:22	-----	-----
SEQ ID NO:23	-KSCCGSRKKGKSS-KK-----	YNYEKRRGINR--SDSNAPLFNMEDIDEGFEG--Y
SEQ ID NO:24	LCC--GLRKKSK---T-----	KAKDKKT--NTKETSQIHALENVDEGVIVPV-
SEQ ID NO:25	-CCCGSRKKSKKKGEKKGLLGGLLYGKKKMMGKNYVKKSAPVFDLEEIEEGLEG--Y	
SEQ ID NO:26	-----NPDKLEKKESETEALTTSDFPDLDVTQLPKRFGNSTLL-----	AESIP
SEQ ID NO:27	-CPCFGRRRKNKK-----	FSKNDMNGDVAALGG
SEQ ID NO:28	-SLCGSRKKSSKSSKK-----	GSDKKKSGKH--VDSTVPVFNLEDIEEGVEGAGF
SEQ ID NO:29	LCF--GSRKNRK---AKT-----	VAADKKK--KNREASKQIHALENIEEGRGHKV-
	901	960
SEQ ID NO:2	-----	-----
SEQ ID NO:4	AENEKASIVNQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCYEDKT	
SEQ ID NO:6	-----	ET
SEQ ID NO:8	ADIEKAGIVNQKLEKKFGQSSVFVASTLLENGGTLKSASPASLLKEAIHVISCYEDKT	
SEQ ID NO:10	AENEKAGIVNQKLEKKFGQSSVFVTSTLLENGGTLKSASPASLLKEAIHVISCYEDKT	
SEQ ID NO:12	-----	-----
SEQ ID NO:14	MDDDKVLMQSMNFEKKFGQSSIFVTSTLMEEGGVPPSSSPAALLKEAIHVISCYEDKT	
SEQ ID NO:16	DD-ERTLLMSQKSLEKRFQSPVFIAATFMEQGGIPPTNPATLLKEAIHVISCYEDKT	
SEQ ID NO:18	TNNEKTSNLTQTKLEKRFQSPVFVASTLLDDGGVPHGVSPASLLKEAIQVISCYEDKT	
SEQ ID NO:20	DD-EKSVLMQSMLEKRFQSAFVASTLMEYGGVPQSSTPESLLKEAIHVISCYEDKS	
SEQ ID NO:22	-----	-----
SEQ ID NO:23	DD-ERSILMSQRSVEKRFQSPVFIAATFMEQGGIPPTNPATLLKEAIHVISCYEDKT	
SEQ ID NO:24	SNVEKRSEATQKLEKKFGQSPVFVASAVLQNGGVPRNASPACLLREAIQVISCYEDKT	
SEQ ID NO:25	EELKSTLMSQKNFEKRFQSPVFIASTLMENGGLPEGTNSTSLIKEAIHVISCYEEKT	
SEQ ID NO:26	AEFQGRPLADHPAV--KYGRPP---GALR---VPRDPLDATTVAESVSVISCWYEDKT	
SEQ ID NO:27	AEGDKELMFEMNFEKTFQSSIFVTSTLMEEGGVPPSSPAVLLKEAIHVISCYEDKT	
SEQ ID NO:28	DD-EKSLMSQMSLEKRFQSAFVASTLMENGGVQSATPETLLKEAIHVISCYEDKT	
SEQ ID NO:29	LNVEQSTEAMQMKLQKKYQSPVFVASARLENGGMARNASPACLLKEAIQVISRGYEDKT	
	961	1020
SEQ ID NO:2	-----HEDILTGFKMARGWISYICMPPRPCFKGSAPINLSDRLNQVLRWAL	
SEQ ID NO:4	GWGKDIGWIYGSVTEDILTGFKMCHGWRSYICIPKRAAFKGSAPINLSDRLHQVLRWAL	
SEQ ID NO:6	EWGKEIGWIYGSVTEDILTGFKMCHGWRSYICIPKRAAFKGSAPINLSDRLHQVLRWAL	
SEQ ID NO:8	DWGKEIGWIYGSITDILTGFKMCHGWRSYICIPKRAAFKGSAPINLSDRLHQVLRWAL	
SEQ ID NO:10	DWGKEIGWIYGSVTEDILTGFKMCHGWRSYICIPKRVAFKGSAPINLSDRLHQVLRWAL	
SEQ ID NO:12	-----WRP-----RRSL	
SEQ ID NO:14	EWGLELGIWYGSITDILTGFKMCHGWRSYICMPKRAAFKGTAPINLSDRLNQVLRWAL	
SEQ ID NO:16	EWGKEIGWIYGSVTEDILTGFKMARGWISYICMPPRPAFKGSAPINLSDRLNQVLRWAL	
SEQ ID NO:18	EWGKEVGWIYGSVTEDILTGFKMCHGWRSYICIPKRAAFKGSAPINLSDRLHQVLRWAL	
SEQ ID NO:20	EWGTEIGWIYGSVTEDILTGFKMARGWRSYICMPKRAAFKGSAPINLSDRLNQVLRWAL	
SEQ ID NO:22	-----	-----
SEQ ID NO:23	EWGKEIGWIYGSVTEDILTGFKMARGWISYICNPPRPAFKGSAPINLSDRLNQVLRWAL	
SEQ ID NO:24	EWGKEIGWIYGSVTEDILTGFKMCHGWRSYICMPKRAAFKGSAPINLSDRLHQVLRWAL	
SEQ ID NO:25	EWGKEIGWIYGSVTEDILTGFKMCHGWRSYICVCPKRAAFKGSAPINLSDRLHQVLRWAL	
SEQ ID NO:26	EWGDRVGWIYGSVTEDVVTGYRMHNRGWSYICITKRDSFRGSAPINLTDRLHQVLRWAT	

Figure 1 (cont'd.)

SEQ ID NO:27 EWGTELGWYGSITEDILTGFKMHCGRWSIYCMKRPAPFKGSAPINLSDRLNQVLRWAL  
 SEQ ID NO:28 DWGSEIGWYGSVTEIDILTGFKMHARGWSIYCMKRPAPFKGSAPINLSDRLNQVLRWAL  
 SEQ ID NO:29 EWGKEIGWYGSVTEIDLTGSKMHSWGHRHVYCTPKLAAPFKGSAPINLSDRLHQVLRWAL

1021 1080  
 SEQ ID NO:2 GSVEILFSRHCPIWYNYGG-RLKLLERMAYINTIVYPITSLPLIAYCVLPAICLLTNKFI  
 SEQ ID NO:4 GSIEIFFSNHCPLWYGYGGG-LKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFI  
 SEQ ID NO:6 GSVEIFMSRHCPLWYAYGG-RLKWLERFAYTNTIVYPFTSIPLLAYCTIPAVCLLTGKFI  
 SEQ ID NO:8 GSVEIFFSKHCPLWYGYGGG-LKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFI  
 SEQ ID NO:10 GSIEIFFSNHCPLWYGYGGG-LKFLERFSYINSIVYPWTSIPLLAYCTLPAICLLTGKFI  
 SEQ ID NO:12 G-----CRL-----  
 SEQ ID NO:14 GSIEIFFSHHCPLWYGFKEKKLKWLERFAYANTTVYPFTSIPLVAYCILPAVCLLTOKFI  
 SEQ ID NO:16 GSIEIFLSRHCPLWYGYNG-KLKPLMRLAYINTIVYPFTSIPLIAYCTLPFAVCLLTNKFI  
 SEQ ID NO:18 GSVEIFFSRHCPIWYGYGGG-LKLLERFSYINSVVPWTSIPLLVYCTLPAICLLTGKFI  
 SEQ ID NO:20 GSVEILFSRHCPLWYGYGG-RLKWLERFAYINTTIYPLTSLPLLVYCTLPAICLLTGKFI  
 SEQ ID NO:22 -----  
 SEQ ID NO:23 GSIEILSRHCPIWYGYHG-RLRLLERIAINTIVYPITSIPLIAYCILPAFCLITDRFI  
 SEQ ID NO:24 GSVEIFLSRHCPIWYGYGGG-LKWLERFSYINSVVPWTSIPLIVYCSLPAVCLLTGKFI  
 SEQ ID NO:25 GSVEIFLSRHCPLWYGYGG-KLKPLMRLAYINTIVYPFTSIPLLAYCTIPAVCLLTGKFI  
 SEQ ID NO:26 GSVEIFFSRNNAI---LASKRLKFLQRLAYLNVGIYPFTSLFLILYCFLPFAFLSFGQFI  
 SEQ ID NO:27 GSVEIFFSRHSPLWYGYKGGKLKWLERFAYANTTIYPFTSIPLLAYCILPAICLLTDKFI  
 SEQ ID NO:28 GSVEILFSRHCPIWYGYSG-RLKWLERFAYVNTTIYPVTAIPLLMYCTLPVAVCLLTNKFI  
 SEQ ID NO:29 GSVEIFLSRHCPIWYGYGGG-LKWLERLSYINSVVPWTSIPLIVYCSLPAICLLTGKFI

1081 1140  
 SEQ ID NO:2 IPEISNYAGMFFILMFASIFATGILELRWSGVGIEDDWRNEQFWVIGGSAHLFAVFQGL  
 SEQ ID NO:4 TPELNNVASLWFMFLFICIFATSILEMRWSGVGIDDWRNEQFWVIGGVSSHFAVFQGL  
 SEQ ID NO:6 IPTLNNLASIWFIALLSIIATSVLELRWSGVSIEDDWRNEQFWVIGGVSAHLFAVFQGL  
 SEQ ID NO:8 TPELTNVASIWFMALFICISVTGILEMRWSGVAIIDDWRNEQFWVIGGVSAHLFAVFQGL  
 SEQ ID NO:10 TPELNNVASLWFMFLFICIFATSILEMRWSGVGIDDWRNEQFWVIGGVSSHFAVFQGL  
 SEQ ID NO:12 -----  
 SEQ ID NO:14 MPPISTFAGLYFVALFSSIIATGILELKWSGVSIIEEWRNEQFWVIGGVSAHLFAVIQGL  
 SEQ ID NO:16 IPEISNFASMWFIILFVSIFTTSILELRWSGVSIEDDWRNEQFWVIGGSAHLFAVFQGL  
 SEQ ID NO:18 VPEISNYASLVFMALFISIAATGILEMQGGVSIIDDWRNEQFWVIGGVSSHFAVFQGL  
 SEQ ID NO:20 MPEISNLASIWFIALLSIFATGILEMRWSGVGIDEWWRNEQFWVIGGISAHLFAVFQGL  
 SEQ ID NO:22 -----  
 SEQ ID NO:23 IPEISNYASIWFIILFISIAVTGILELRWSGVSIEDDWRNEQFWVIGGSAHLFAVFQGL  
 SEQ ID NO:24 VPEISNYAGILFMLMFISIAVTGILEMQGGVSIIDDWRNEQFWVIGGASSHLFALFQGL  
 SEQ ID NO:25 IPTLSNLTSVWFLALFLSIIATGVLELRWSGVSIQDWRNEQFWVIGGVSAHLFAVFQGL  
 SEQ ID NO:26 VRTLSISFLVYLLMITICLIGLAVLEVKWSGIGLEEWRNEQWWLISGTSSHLVAVVQGV  
 SEQ ID NO:27 MPPISTFASLFFISLFSIIVTGILELRWSGVSIIEEWRNEQFWVIGGISAHLFAVFQGL  
 SEQ ID NO:28 IPQISNLASIWFIISLFSIFATGILKMKWNGVGDQWRNEQFWVIGGVSAHLFAVFQGL  
 SEQ ID NO:29 VPEISNYASILFMALFSSIAITGILEMQGKVGIDDWRNEQFWVIGGVSAHLFALFQGL

1141 1200  
 SEQ ID NO:2 LKVLAGIDTNFTVTSKANDGEDG--FAELYVFKWTSLLIPPTTVLVINLVGMVAGISYAI  
 SEQ ID NO:4 LKVIAGVDTSTFTVTSKGGDD--EE-FSELYTFKWTTLLIPPTTLLLNFIGVAGISNAI  
 SEQ ID NO:6 LKVLGGVDTSTFTVTSKAAGDEADA-FGDLYLFKWTTLLVPPTTLIIINMVGIVAGVSDAV  
 SEQ ID NO:8 LKVFAIDTSTFTVTSKAGDD--EE-FSELYTFKWTTLLIPPTTLLLNFIGVAGISNAI  
 SEQ ID NO:10 LKVIAGVDTSTFTVTSKGGDD--EE-FSELYTFKWTTLLIPPTTLLLNFIGVAGVSNAI  
 SEQ ID NO:12 -----  
 SEQ ID NO:14 LKVLAGIDTNFTVTSKATDDE-E--FGELYTFKWTTLLIPPTTILIIINIVGVVAGISDAI  
 SEQ ID NO:16 LKVLAGIDTNFTVTSKASDEGD--FAELYVFKWTSLLIPPTTVLVINLVGIVAGVSYAI  
 SEQ ID NO:18 LKVLAVNTNFTVTSKAADD--GE-FSELYIFKWTSLLIPPTTLLIMNIVGVVAGISDAI  
 SEQ ID NO:20 LKVLAGIDTNFTVTSKANDEGD--FAELYMFKWTTLLIPPTTILIIINMVGIVAGTSYAI  
 SEQ ID NO:22 -----  
 SEQ ID NO:23 LKVLAGIDTNFTVTSKATDEGD--FAELYIFKWTTALLIPPTTVLLVNLIGIVAGVSYAV



Figure 1 (cont'd.)

```

SEQ ID NO:24 LKVLAVNTNFTVTSKAADD--GA-FSELYIFKWTTLLIPPTLLIINIIGVIVGVSDAI
SEQ ID NO:25 LKVLAVDTNFTVTAKAADDTE---FGELYLFKWTTLLIPPTLLIILNMVGVVAGVSDAI
SEQ ID NO:26 LKVIAGIEISFTLTTSKGGDDNEDIYADLYIVKWSSLMIPPVIVAMVNIIAIVVAFIRTI
SEQ ID NO:27 LKILAGIDTNFTVTSKATDDD-D--FGELYAFKWTTLLIPPTVLIINIVGVVAGISDAI
SEQ ID NO:28 LKVLAVDTNFTVTSKASDEDDGD--FAELYMFKWTTLLIPPTLLIINLVGVVAGISYVI
SEQ ID NO:29 LKVLAVDTNFTVTSKAADD--GE-FSDLYLFKWTSLLIPPMTLLIINIVGVIVGVSDAI

```

```

1201
SEQ ID NO:2 NSGYQSWGPLFGKLFFSIWVILHLYPFLKGLMGKQNRTPPTIVIVWSILLASIFSLLWVKI
SEQ ID NO:4 NNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPPTIVIVWSILLASIFSLLWVRI
SEQ ID NO:6 NNGYGSWGPLFGKLFFSFWVIVHLYPFLKGLMGRQNRTPPTIVVLSILLASIFSLLWVRI
SEQ ID NO:8 NNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPPTIVIVWSILLASIFSLLWVRV
SEQ ID NO:10 NNGYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPPTIVIVWSILLASIFSLLWVRI
SEQ ID NO:12 -----LGEDER-----LWSRM
SEQ ID NO:14 NNGYQSWGPLFGKLFFSFWVIVHLYPFLKGLMGRQNRTPPTIVVWSVLLASIFSLLWVRI
SEQ ID NO:16 NSGYQSWGPLFGKLFFAIWVIAHLYPFLKGLLGRQNRTPPTIVIVWSVLLASIFSLLWVRI
SEQ ID NO:18 NNGYDSWGPLFGRLFFALWVILHLYPFLKGLLGKQDRMPTIILVWSILLASILTLMWVRI
SEQ ID NO:20 NSGYQSWGPLFGKLFFAFWVIVHLYPFLKGLMGRQNRTPPTIVIVWAVLLASIFSLLWVRV
SEQ ID NO:22 -----
SEQ ID NO:23 NSGYQSWGPLFGKLFFALWVIAHLYPFLKGLLGRQNRTPPTIVIVWSVLLASIFSLLWVRI
SEQ ID NO:24 SNGYDSWGPLFGRLFFALWVIVHLYPFLKGLMLGKQDKMPTIIVVWSILLASILTLLWVRV
SEQ ID NO:25 NNGYGSWGPLFGKLFFAFWVILHLYPFLKGLMGRQNRTPPTIVVLSILLASIFSLLWVRI
SEQ ID NO:26 YQAVPQWSKLIGGAFFSFWVLAHLYPFAKGLMGRRGKTPTIVFVWAGLIAITISLLWTAI
SEQ ID NO:27 NNGYQSWGPLFGKLFFSFWVIVHLYPFLKGLMGRQNRTPPTIVVWSVLLASIFSLLWVRI
SEQ ID NO:28 NSGYQSWGPLFGKLFFAFWVIIHLYPFLKGLMGRQNRTPPTIVVWSILLASIFSLLWVRI
SEQ ID NO:29 SNGYDSWGPLFGRLFFALWVIIHLYPFLKGLLGKQDRMPTIIVVWSILLASILTLLWVRV

```

```

1261 1284
SEQ ID NO:2 DPFISDTQKAVAM-GQCGVNC---
SEQ ID NO:4 DPFLAKDDGPLL--EECGLDCN--
SEQ ID NO:6 DPFIPKAKGPILKP-C-GVEC---
SEQ ID NO:8 DPFLAKSNGPLL--EECGLDCN--
SEQ ID NO:10 DPFLAKDDGPLL--EECGLDCN--
SEQ ID NO:12 KQMVILSGPR-----
SEQ ID NO:14 DPFVLKTKGPDTKL--CGINC---
SEQ ID NO:16 DPFTSDSNKLT--NGQCGINC---
SEQ ID NO:18 NPFVSRD-GPVL--EICGLNCDES
SEQ ID NO:20 DPFTTRLAGPNI--QTCGINC---
SEQ ID NO:22 -----
SEQ ID NO:23 NPFVDANPNANNFNGKGGVF----
SEQ ID NO:24 NPFVAKG-GPVL--EICGLNCGN-
SEQ ID NO:25 DPFLPKQTGPVLKQ-C-GVEC---
SEQ ID NO:26 NP---NTGPAAAAEGVGGGGFQFP
SEQ ID NO:27 DPFVLKTKGPDTSK--CGINC---
SEQ ID NO:28 DPFTTRVTGPDV--EQCGINC---
SEQ ID NO:29 NPFVAKG-GPIL--EICGLDC--L

```

SEQUENCE LISTING

&lt;110&gt; E. I. du Pont de Nemours and Company

&lt;120&gt; Plant Cellulose Synthases

&lt;130&gt; BB-1170

&lt;140&gt;

&lt;141&gt;

&lt;150&gt; 60/092,844

&lt;151&gt; July 14, 1998

&lt;160&gt; 29

&lt;170&gt; Microsoft Office 97

&lt;210&gt; 1

&lt;211&gt; 1221

&lt;212&gt; DNA

&lt;213&gt; Hordeum vulgare

&lt;400&gt; 1

```

gcacgaggat attcttactg ggtttaaaat gcacgcaaga ggttggatat caatctactg   60
catgccacca cgacctgtt tcaaggggtc tgcgccaatc aatctctctg accgtctcaa  120
tcaagttctc cgggtgggctc ttgggtcagt tgaaattctg tttagcagac attgtcctat  180
ctgggtacaat tacgggtgggc gggtgaaaact tctggagagg atggcttaca tcaacaccat  240
tgtttatcca ataacatccc ttccacttat cgcctattgt gtgcttcctg ctatctgtct  300
cctcaccaac aaatttatca ttcccgagat cagtaactat gctgggatgt tctttattct  360
tatgtttgcc tccatctttg ccacgggtat attggagctg cgatggagtg gtgtcggcat  420
cgaggactgg tggagaaacg agcagttctg ggttattggg ggcacatctg cccatctttt  480
cgcagtgttc caggggtctg tgaaggtgtt ggccgggatt gacaccaact tcacggttac  540
ctggaaggca aacgacgagg atggcgattt tgctgagtta tacgtgttca agtggaccag  600
tctcctcatt cctccgacca ccgtccttgt gattaacctg gtgggcatgg tggcaggcat  660
atcatatgcc atcaacagcg gttaccagtc ttgggggtcca ctcttcggaa agctcttctt  720
ctcaatctgg gtgatcctcc atctctaccc ctctctcaag ggtctcatgg ggaagcagaa  780
ccgcacgcca accatcgtca ttgtttggtc catcctccta gcctccatct tctccctccc  840
gtgggtgaag atcgaccctt tcatatccga taccagaaa gccgtcgcca tggggcagtg  900
tggcgtaaac tgctgatcgg cgccgaagag tatctgcccc cctcgtgtaa ataccggagg  960
gggttggatg ggattttgtt gttgtagatg aagacggagt tttatgtaag ttattattgc 1020
cccttcgtgc tgagaagcac aaaccgtgaa ccctacgaaa cctgcagcgt acattgtgat 1080
ttttttctcc ttttcttttc atctgtgata cctgttgttt cttcttagag tatattatgt 1140
cagaacgtat ctatagttct atacacacta tgacaccaac tatttatata aggcagctgt 1200
tgcatcaact cttctgcaaa a                                     1221

```

&lt;210&gt; 2

&lt;211&gt; 304

&lt;212&gt; PRT

&lt;213&gt; Hordeum vulgare

&lt;400&gt; 2

```

His Glu Asp Ile Leu Thr Gly Phe Lys Met His Ala Arg Gly Trp Ile
  1             5             10             15
Ser Ile Tyr Cys Met Pro Pro Arg Pro Cys Phe Lys Gly Ser Ala Pro
          20             25             30
Ile Asn Leu Ser Asp Arg Leu Asn Gln Val Leu Arg Trp Ala Leu Gly
  35             40             45

```

Ser Val Glu Ile Leu Phe Ser Arg His Cys Pro Ile Trp Tyr Asn Tyr  
 50 55 60  
 Gly Gly Arg Leu Lys Leu Leu Glu Arg Met Ala Tyr Ile Asn Thr Ile  
 65 70 75 80  
 Val Tyr Pro Ile Thr Ser Leu Pro Leu Ile Ala Tyr Cys Val Leu Pro  
 85 90 95  
 Ala Ile Cys Leu Leu Thr Asn Lys Phe Ile Ile Pro Glu Ile Ser Asn  
 100 105 110  
 Tyr Ala Gly Met Phe Phe Ile Leu Met Phe Ala Ser Ile Phe Ala Thr  
 115 120 125  
 Gly Ile Leu Glu Leu Arg Trp Ser Gly Val Gly Ile Glu Asp Trp Trp  
 130 135 140  
 Arg Asn Glu Gln Phe Trp Val Ile Gly Gly Thr Ser Ala His Leu Phe  
 145 150 155 160  
 Ala Val Phe Gln Gly Leu Leu Lys Val Leu Ala Gly Ile Asp Thr Asn  
 165 170 175  
 Phe Thr Val Thr Ser Lys Ala Asn Asp Glu Asp Gly Asp Phe Ala Glu  
 180 185 190  
 Leu Tyr Val Phe Lys Trp Thr Ser Leu Leu Ile Pro Pro Thr Thr Val  
 195 200 205  
 Leu Val Ile Asn Leu Val Gly Met Val Ala Gly Ile Ser Tyr Ala Ile  
 210 215 220  
 Asn Ser Gly Tyr Gln Ser Trp Gly Pro Leu Phe Gly Lys Leu Phe Phe  
 225 230 235 240  
 Ser Ile Trp Val Ile Leu His Leu Tyr Pro Phe Leu Lys Gly Leu Met  
 245 250 255  
 Gly Lys Gln Asn Arg Thr Pro Thr Ile Val Ile Val Trp Ser Ile Leu  
 260 265 270  
 Leu Ala Ser Ile Phe Ser Leu Leu Trp Val Lys Ile Asp Pro Phe Ile  
 275 280 285  
 Ser Asp Thr Gln Lys Ala Val Ala Met Gly Gln Cys Gly Val Asn Cys  
 290 295 300

<210> 3  
 <211> 3776  
 <212> DNA  
 <213> Zea mays

<400> 3  
 gcgcgcgcgc caggcgcaac gcaacaaagg gaaaccccag ccggaggagc aaaagctagc 60  
 aagcgtgtcc ctccccctcc ctcaactccg ttctattcca ttcccccca gacgccgcta 120  
 ccgccgccgc cgcacgcacg cttgccccgg gatctggaga tctggtagcg ccagggggat 180  
 ggaggccagc gccgggctgg tcgcccggctc gcacaaccgg aacgagctcg tcgtcatccg 240  
 ccgcgatggc gagccagggc cgaagcccat ggaccagcgg aacggccagg tgtgccagat 300

ttgcggcgac	gacgtggggc	gcaaccccg	cgaggagccg	ttcgtggcct	gcaacgagt	360
cgccttcccc	atctgccggg	actgctacga	gtacgagcgc	cgcgagggca	cgcagaact	420
ccccagtg	aagaccgc	tcaagcgct	caaggggtgc	gcgcgcgtgc	ccggggacga	480
ggaggaggac	ggcgtcgacg	acctggagaa	cgagttcaac	tggagcgaca	agcacgactc	540
ccagtacctc	gccgagtc	tgctccacgc	ccacatgagc	tacggccgcg	gcgcccac	600
cgacggcg	ccgagccat	tccaccccat	ccccaatgtt	cccctcctca	ccaacggaca	660
gatggtcgat	gacatccgc	cggaccagca	cgcccttg	ccctcgttcg	tgggtggcg	720
ggggaagagg	attcaccctc	tcccgtagc	ggatcccaac	cttcctgtgc	aaccgaggtc	780
tatggaccct	tccaaggatc	tcgcccata	tggctacggg	agcgtagcat	ggaaggagag	840
gatggagagc	tggaagcaga	agcaggagag	gatgcaccag	acgaggaacg	atggcgccg	900
cgatgatggt	gatgatgcag	atctaccact	aatggatgaa	gctagacagc	cattgtccag	960
aaagatcccc	cctccttcaa	gccaaatcaa	cccctatagg	atgattataa	taattcggt	1020
agtggttttg	tgtttcttct	tccactaccg	atgcatgc	ccggtgcctg	atgcatttgc	1080
tttatggctc	atatctgtga	tctgtgaaat	ttggtttgcc	atgtcttgg	ttcttgacca	1140
gtttccaaag	tggtttccta	tcgagaggga	aacctatctt	gaccggctga	gtttaagggt	1200
tgacaaggaa	gggcatcctt	ctcaactcgc	ccctgttgat	ttctttgtca	gtacggttga	1260
tcccttgaag	gaacctccat	tggtcactgc	taatactgtt	ctatctatcc	tttcggtgga	1320
ttatccagtt	gataaggttt	catgctacgt	ttctgatgat	ggtgctgcca	tgctgacatt	1380
tgaagcattg	tctgaaacat	ctgaatttgc	aaaggaaagg	gttcctttct	gcaaaagata	1440
tagccttgag	cctcgtgctc	cagagtggt	cttccaacag	aagatagact	acctgaaaga	1500
caaggtggcg	ccaaactttg	ttagagaacg	gagagcaatg	aagagagagt	atgaggaaat	1560
caaggtcaga	atcaatgcct	tggttgctaa	agcccaaaag	gttcctgagg	aaggatggac	1620
aatgcaggat	ggaactccat	ggcccggaaa	taatgtccgt	gatcatcctg	gaatgattca	1680
ggttttccct	ggtcaaagt	gtggccatga	tgtggaagga	aatgagctgc	ctcgatttgt	1740
ttatgtttca	agagaaaaac	ggccaggcta	caaccatcac	aagaaggctg	gtgctatgaa	1800
tgcatgtgtc	cgagtctctg	ctgtactaac	taatgtcctc	tatttgctga	acttggattg	1860
tgatcactat	atcaataata	gtaaggctat	aaaggaaagca	atgtgtttta	tgatggatcc	1920
tttgcttgg	aagaaagt	gctatgtgca	gtttcctcaa	agatttgatg	ggattgatcg	1980
ccatgatcga	tatgctaaca	gaaatgttgt	ctttttcgat	atcaacatga	aaggtttgg	2040
tggtatccag	ggcccaattt	atgtgggtac	tggatgtgtc	ttcagaaggc	aggcattata	2100
tggtacgat	gctcccaaaa	caaagaagcc	accatcaaga	acttgcaact	gctggccaaa	2160
gtggtgcatt	tgctgttgct	gttttggtta	caggaagacc	aagaagaaga	ccaagacctc	2220
taaaccataa	tttgagaaga	taaagaaact	ttttaagaaa	aaggaaaatc	aagcccctgc	2280
atatgctctt	ggtgaaattg	atgaagccgc	tccaggagct	gaaaatgaaa	aggctagat	2340
tgtaaatcaa	cagaagttgg	aaaagaaatt	tggccagtct	tcagtttttg	ttgcatccac	2400
acttcttgag	aatggtggaa	ccctgaagag	tgccagtcca	gcttctcttc	tgaaggaaag	2460
tatacatgtc	atcagttgtg	gatatgaaga	caaaacaggc	tggggaaaag	atattgggtg	2520
gatttatgga	tcagtcacag	aagatattct	tactgggttt	aagatgcact	gccatgggtg	2580
gcggtcaatt	tactgcatac	ctaaacgggc	cgcttcaaaa	ggttccgcac	ctctcaatct	2640
ttccgatcgt	tctcaccagg	ttcttcggtg	ggctcttggt	tcaattgaaa	ttttcttcag	2700
caaccactgc	cctctctggt	atgggtatgg	tgggtgacta	aagttcctgg	aaaggttttc	2760
gtacattaac	tccatcgtat	acccttggac	atctatcccg	ctcttgccct	attgcacatt	2820
gcctgccatc	tgcttgctga	cagggaatt	tatcacgcca	gagcttaaca	atgttgccag	2880
cctctgggtc	atgtcacttt	tcatctgcat	ttttgctacg	agcatcctgg	aaatgagatg	2940
gagtggtgta	ggcatcgatg	actggtggag	aaacgagcag	ttttgggtca	ttggaggcgt	3000
gtcttcacat	ctctttgctg	tgttccaggg	actcctcaag	gtcatagctg	gtgtagacac	3060
gagcttcact	gtgacatcca	agggcgagga	cgacaggag	ttctcagagc	tgtacacatt	3120
caaattggacg	acccttctga	tacctccgac	aacctgctc	ctactgaact	tcattggagt	3180
ggtagctggc	atctccaatg	cgatcaacaa	cggatatgaa	tcattggggcc	ccctgttcgg	3240
gaagctcttc	tttgcatttt	gggtgatcgt	ccatctttac	ccgttccctca	agggctctggt	3300
tgggaggcag	aacaggacgc	caacgattgt	cattgtctgg	tccatcctcc	tggcttcgat	3360
cttctcgctg	ctttgggtcc	ggatcgaccc	gttccttgcg	aaggatgatg	gtcccctgtt	3420
ggaggagtgt	ggtctggatt	gcaactagga	ggtcagcacg	tggacttccc	cgtcagtgtg	3480
tggtcgaaaga	agattttttg	cagatgtttt	gtgcccata	ttcttttttc	aatttttgtc	3540
cctctgtaga	tagaaacaag	gggagaagg	gaaaaaaagt	acttgtattt	ctttgttcc	3600
atggtggtgg	tggtggtggg	cggctcagcc	tcgtgagtg	agtattgggc	aaaccggagg	3660
ctgcggcaac	cttgtgcagt	tcggccacga	atatactagg	gaagatcgcg	accaatcaat	3720
caatcgatga	ccgagttcaa	ttgttcagca	aaaaaaaaaa	aaaaaaaaaa	aaaaaa	3776

<210> 4  
 <211> 1148  
 <212> PRT  
 <213> Zea mays  
  
 <400> 4  
 Arg Ala Ala Gln Ala Gln Arg Asn Lys Gly Lys Pro Gln Pro Glu Glu  
 1 5 10 15  
 Gln Lys Leu Ala Ser Val Ser Leu Pro Leu Pro His Ser Arg Phe Ile  
 20 25 30  
 Pro Phe Pro Pro Arg Arg Arg Tyr Arg Arg Arg Arg Thr His Ala Cys  
 35 40 45  
 Pro Gly Ile Trp Arg Ser Gly Ser Ala Arg Gly Met Glu Ala Ser Ala  
 50 55 60  
 Gly Leu Val Ala Gly Ser His Asn Arg Asn Glu Leu Val Val Ile Arg  
 65 70 75 80  
 Arg Asp Gly Glu Pro Gly Pro Lys Pro Met Asp Gln Arg Asn Gly Gln  
 85 90 95  
 Val Cys Gln Ile Cys Gly Asp Asp Val Gly Arg Asn Pro Asp Gly Glu  
 100 105 110  
 Pro Phe Val Ala Cys Asn Glu Cys Ala Phe Pro Ile Cys Arg Asp Cys  
 115 120 125  
 Tyr Glu Tyr Glu Arg Arg Glu Gly Thr Gln Asn Cys Pro Gln Cys Lys  
 130 135 140  
 Thr Arg Phe Lys Arg Leu Lys Gly Cys Ala Arg Val Pro Gly Asp Glu  
 145 150 155 160  
 Glu Glu Asp Gly Val Asp Asp Leu Glu Asn Glu Phe Asn Trp Ser Asp  
 165 170 175  
 Lys His Asp Ser Gln Tyr Leu Ala Glu Ser Met Leu His Ala His Met  
 180 185 190  
 Ser Tyr Gly Arg Gly Ala Asp Leu Asp Gly Val Pro Gln Pro Phe His  
 195 200 205  
 Pro Ile Pro Asn Val Pro Leu Leu Thr Asn Gly Gln Met Val Asp Asp  
 210 215 220  
 Ile Pro Pro Asp Gln His Ala Leu Val Pro Ser Phe Val Gly Gly Gly  
 225 230 235 240  
 Gly Lys Arg Ile His Pro Leu Pro Tyr Ala Asp Pro Asn Leu Pro Val  
 245 250 255  
 Gln Pro Arg Ser Met Asp Pro Ser Lys Asp Leu Ala Ala Tyr Gly Tyr  
 260 265 270  
 Gly Ser Val Ala Trp Lys Glu Arg Met Glu Ser Trp Lys Gln Lys Gln  
 275 280 285

Glu Arg Met His Gln Thr Arg Asn Asp Gly Gly Gly Asp Asp Gly Asp  
 290 295 300  
 Asp Ala Asp Leu Pro Leu Met Asp Glu Ala Arg Gln Pro Leu Ser Arg  
 305 310 315 320  
 Lys Ile Pro Leu Pro Ser Ser Gln Ile Asn Pro Tyr Arg Met Ile Ile  
 325 330 335  
 Ile Ile Arg Leu Val Val Leu Cys Phe Phe Phe His Tyr Arg Val Met  
 340 345 350  
 His Pro Val Pro Asp Ala Phe Ala Leu Trp Leu Ile Ser Val Ile Cys  
 355 360 365  
 Glu Ile Trp Phe Ala Met Ser Trp Ile Leu Asp Gln Phe Pro Lys Trp  
 370 375 380  
 Phe Pro Ile Glu Arg Glu Thr Tyr Leu Asp Arg Leu Ser Leu Arg Phe  
 385 390 395 400  
 Asp Lys Glu Gly His Pro Ser Gln Leu Ala Pro Val Asp Phe Phe Val  
 405 410 415  
 Ser Thr Val Asp Pro Leu Lys Glu Pro Pro Leu Val Thr Ala Asn Thr  
 420 425 430  
 Val Leu Ser Ile Leu Ser Val Asp Tyr Pro Val Asp Lys Val Ser Cys  
 435 440 445  
 Tyr Val Ser Asp Asp Gly Ala Ala Met Leu Thr Phe Glu Ala Leu Ser  
 450 455 460  
 Glu Thr Ser Glu Phe Ala Lys Lys Trp Val Pro Phe Cys Lys Arg Tyr  
 465 470 475 480  
 Ser Leu Glu Pro Arg Ala Pro Glu Trp Tyr Phe Gln Gln Lys Ile Asp  
 485 490 495  
 Tyr Leu Lys Asp Lys Val Ala Pro Asn Phe Val Arg Glu Arg Arg Ala  
 500 505 510  
 Met Lys Arg Glu Tyr Glu Glu Phe Lys Val Arg Ile Asn Ala Leu Val  
 515 520 525  
 Ala Lys Ala Gln Lys Val Pro Glu Glu Gly Trp Thr Met Gln Asp Gly  
 530 535 540  
 Thr Pro Trp Pro Gly Asn Asn Val Arg Asp His Pro Gly Met Ile Gln  
 545 550 555 560  
 Val Phe Leu Gly Gln Ser Gly Gly His Asp Val Glu Gly Asn Glu Leu  
 565 570 575  
 Pro Arg Leu Val Tyr Val Ser Arg Glu Lys Arg Pro Gly Tyr Asn His  
 580 585 590  
 His Lys Lys Ala Gly Ala Met Asn Ala Leu Val Arg Val Ser Ala Val  
 595 600 605

Leu Thr Asn Ala Pro Tyr Leu Leu Asn Leu Asp Cys Asp His Tyr Ile  
 610 615 620  
 Asn Asn Ser Lys Ala Ile Lys Glu Ala Met Cys Phe Met Met Asp Pro  
 625 630 635 640  
 Leu Leu Gly Lys Lys Val Cys Tyr Val Gln Phe Pro Gln Arg Phe Asp  
 645 650 655  
 Gly Ile Asp Arg His Asp Arg Tyr Ala Asn Arg Asn Val Val Phe Phe  
 660 665 670  
 Asp Ile Asn Met Lys Gly Leu Asp Gly Ile Gln Gly Pro Ile Tyr Val  
 675 680 685  
 Gly Thr Gly Cys Val Phe Arg Arg Gln Ala Leu Tyr Gly Tyr Asp Ala  
 690 695 700  
 Pro Lys Thr Lys Lys Pro Pro Ser Arg Thr Cys Asn Cys Trp Pro Lys  
 705 710 715 720  
 Trp Cys Ile Cys Cys Cys Cys Phe Gly Asn Arg Lys Thr Lys Lys Lys  
 725 730 735  
 Thr Lys Thr Ser Lys Pro Lys Phe Glu Lys Ile Lys Lys Leu Phe Lys  
 740 745 750  
 Lys Lys Glu Asn Gln Ala Pro Ala Tyr Ala Leu Gly Glu Ile Asp Glu  
 755 760 765  
 Ala Ala Pro Gly Ala Glu Asn Glu Lys Ala Ser Ile Val Asn Gln Gln  
 770 775 780  
 Lys Leu Glu Lys Lys Phe Gly Gln Ser Ser Val Phe Val Ala Ser Thr  
 785 790 795 800  
 Leu Leu Glu Asn Gly Gly Thr Leu Lys Ser Ala Ser Pro Ala Ser Leu  
 805 810 815  
 Leu Lys Glu Ala Ile His Val Ile Ser Cys Gly Tyr Glu Asp Lys Thr  
 820 825 830  
 Gly Trp Gly Lys Asp Ile Gly Trp Ile Tyr Gly Ser Val Thr Glu Asp  
 835 840 845  
 Ile Leu Thr Gly Phe Lys Met His Cys His Gly Trp Arg Ser Ile Tyr  
 850 855 860  
 Cys Ile Pro Lys Arg Ala Ala Phe Lys Gly Ser Ala Pro Leu Asn Leu  
 865 870 875 880  
 Ser Asp Arg Leu His Gln Val Leu Arg Trp Ala Leu Gly Ser Ile Glu  
 885 890 895  
 Ile Phe Phe Ser Asn His Cys Pro Leu Trp Tyr Gly Tyr Gly Gly Gly  
 900 905 910  
 Leu Lys Phe Leu Glu Arg Phe Ser Tyr Ile Asn Ser Ile Val Tyr Pro  
 915 920 925

Trp Thr Ser Ile Pro Leu Leu Ala Tyr Cys Thr Leu Pro Ala Ile Cys  
 930 935 940  
 Leu Leu Thr Gly Lys Phe Ile Thr Pro Glu Leu Asn Asn Val Ala Ser  
 945 950 955 960  
 Leu Trp Phe Met Ser Leu Phe Ile Cys Ile Phe Ala Thr Ser Ile Leu  
 965 970 975  
 Glu Met Arg Trp Ser Gly Val Gly Ile Asp Asp Trp Trp Arg Asn Glu  
 980 985 990  
 Gln Phe Trp Val Ile Gly Gly Val Ser Ser His Leu Phe Ala Val Phe  
 995 1000 1005  
 Gln Gly Leu Leu Lys Val Ile Ala Gly Val Asp Thr Ser Phe Thr Val  
 1010 1015 1020  
 Thr Ser Lys Gly Gly Asp Asp Glu Glu Phe Ser Glu Leu Tyr Thr Phe  
 1025 1030 1035 1040  
 Lys Trp Thr Thr Leu Leu Ile Pro Pro Thr Thr Leu Leu Leu Leu Asn  
 1045 1050 1055  
 Phe Ile Gly Val Val Ala Gly Ile Ser Asn Ala Ile Asn Asn Gly Tyr  
 1060 1065 1070  
 Glu Ser Trp Gly Pro Leu Phe Gly Lys Leu Phe Phe Ala Phe Trp Val  
 1075 1080 1085  
 Ile Val His Leu Tyr Pro Phe Leu Lys Gly Leu Val Gly Arg Gln Asn  
 1090 1095 1100  
 Arg Thr Pro Thr Ile Val Ile Val Trp Ser Ile Leu Leu Ala Ser Ile  
 1105 1110 1115 1120  
 Phe Ser Leu Leu Trp Val Arg Ile Asp Pro Phe Leu Ala Lys Asp Asp  
 1125 1130 1135  
 Gly Pro Leu Leu Glu Glu Cys Gly Leu Asp Cys Asn  
 1140 1145

&lt;210&gt; 5

&lt;211&gt; 1189

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 5

gcacgagacc	gagtggggca	aggagattgg	gtggatctat	gggtcgggtga	cagaggatat	60
cctgacgggg	ttcaagatgc	actgccgggg	gtggaagtcc	gtgtactgca	cgccgacacg	120
gccggcggttc	aagggggtcgg	cgcccatcaa	cttgtctgat	cgtctccacc	aggtgctgcg	180
ctgggcgctg	gggtccgtgg	agatcttcat	gagccgccac	tgcccgtctt	ggtacgccta	240
cgggcgccgg	ctcaagtggc	tggagcgctt	cgcctacacc	aacaccatcg	tgtaccctt	300
cacctccatc	cgctcctcgg	cctactgcac	catccccgcc	gtctgcctgc	tcaccggcaa	360
gttcatcatt	cccacgctga	acaacctcgc	cagcatctgg	ttcatcgcg	tcttcctgtc	420
catcatcgcg	acgagcgctc	tggagctgcg	gtggagcggg	gtgagcatcg	aggactgggt	480
gcgcaacgag	cagttctggg	tcatcggcgg	cgtgtccgcg	catctcttcg	ccgtgttcca	540
gggcttcttc	aaggttctgg	gcggcggtga	caccagcttc	accgtcacct	ccaaggcggc	600
cgggcgacgag	gccgacgcct	tcggggacct	ctacctcttc	aagtggacca	ccctgctggt	660
gccccccacc	acgctcatca	tcatcaacat	ggtgggcata	gtggccggcg	tgtccgacgc	720



```

cgtcaacaac ggctacggct cctggggccc gctcttcggc aagctcttct tctccttctg 780
ggtcacgcgc cacctctacc cgttcctcaa ggggctcatg gggaggcaga accggacgcc 840
caccatcgtc gtgctctggg ccatcctcct cgcctccatc ttctcgctcg tctgggtcag 900
gatcgaccgg tttatcccga aggccaaagg ccccatcctc aagccatgcg gagtcgagtg 960
ctgagctcac ctagctacct tcttggtgca tgtacggacg ccgccgtgcg ttgggacata 1020
caggcacttt tgggccaggc tactcatgtt cgactttttt ttttaatttg tacaagattt 1080
gtgatcgagt gactgagtga gacagagtgt tgggtgtaag aactgtgatg gaattcactc 1140
aaattaatgg acattttttt tcttcaactg caaaaaaaaa aaaaaaaaaa 1189

```

<210> 6  
 <211> 320  
 <212> PRT  
 <213> Zea mays

```

<400> 6
His Glu Thr Glu Trp Gly Lys Glu Ile Gly Trp Ile Tyr Gly Ser Val
  1              5              10              15

Thr Glu Asp Ile Leu Thr Gly Phe Lys Met His Cys Arg Gly Trp Lys
      20              25              30

Ser Val Tyr Cys Thr Pro Thr Arg Pro Ala Phe Lys Gly Ser Ala Pro
      35              40              45

Ile Asn Leu Ser Asp Arg Leu His Gln Val Leu Arg Trp Ala Leu Gly
  50              55              60

Ser Val Glu Ile Phe Met Ser Arg His Cys Pro Leu Trp Tyr Ala Tyr
  65              70              75              80

Gly Gly Arg Leu Lys Trp Leu Glu Arg Phe Ala Tyr Thr Asn Thr Ile
      85              90              95

Val Tyr Pro Phe Thr Ser Ile Pro Leu Leu Ala Tyr Cys Thr Ile Pro
      100             105             110

Ala Val Cys Leu Leu Thr Gly Lys Phe Ile Ile Pro Thr Leu Asn Asn
      115             120             125

Leu Ala Ser Ile Trp Phe Ile Ala Leu Phe Leu Ser Ile Ile Ala Thr
      130             135             140

Ser Val Leu Glu Leu Arg Trp Ser Gly Val Ser Ile Glu Asp Trp Trp
      145             150             155             160

Arg Asn Glu Gln Phe Trp Val Ile Gly Gly Val Ser Ala His Leu Phe
      165             170             175

Ala Val Phe Gln Gly Phe Leu Lys Val Leu Gly Gly Val Asp Thr Ser
      180             185             190

Phe Thr Val Thr Ser Lys Ala Ala Gly Asp Glu Ala Asp Ala Phe Gly
      195             200             205

Asp Leu Tyr Leu Phe Lys Trp Thr Thr Leu Leu Val Pro Pro Thr Thr
      210             215             220

Leu Ile Ile Ile Asn Met Val Gly Ile Val Ala Gly Val Ser Asp Ala
      225             230             235             240

```

Val Asn Asn Gly Tyr Gly Ser Trp Gly Pro Leu Phe Gly Lys Leu Phe  
 245 250 255

Phe Ser Phe Trp Val Ile Val His Leu Tyr Pro Phe Leu Lys Gly Leu  
 260 265 270

Met Gly Arg Gln Asn Arg Thr Pro Thr Ile Val Val Leu Trp Ser Ile  
 275 280 285

Leu Leu Ala Ser Ile Phe Ser Leu Val Trp Val Arg Ile Asp Pro Phe  
 290 295 300

Ile Pro Lys Ala Lys Gly Pro Ile Leu Lys Pro Cys Gly Val Glu Cys  
 305 310 315 320

<210> 7  
 <211> 3786  
 <212> DNA  
 <213> Zea mays

<400> 7  
 ccacagctca tataccaaga gccggagcag cttagcgag cccagagcgg cgccgcgcc 60  
 agcacaaccc ccaccgcga cagccgcgtg cgcattgtg cggtcgccgc ggccgggaga 120  
 ccagaggagg ggaggactac gtgcatttcg ctgtgccgcc gccgcggggt tegtgcgcga 180  
 gcgagatccg gcggggaggc gccggggggc tgagatggag gctagcgagg ggctggtggc 240  
 cggctcgcat aaccggaacg agctggtggt gatccgcgc gaccgcgagt cgggagccgc 300  
 gggcgggcggc gcggcgccgc gggcgaggc gccgtgccag atatgcggcg acgaggtcgg 360  
 ggtgggcttc gacggggagc cttcgtggc gtgcaacgag tgcgccttc ccgtctgccg 420  
 cgctgtctac gactacgagc gccgcgagg ctcgcaagcg tgcccgcagt gcaggaccgc 480  
 ctacaagcgc ctcaagggtt gccgcgggt gccggcgac gaggaggagg acggcgctga 540  
 cgacctggag ggcgagttcg gcctgcagga cggcgccgc cacgaggacg accgcagta 600  
 cgtcgccgag tccatgtca gggcgagat gagctacggc cgcggcgagg acgcgcaccc 660  
 cggcttcagc ccgctcccca acgtgccgt cctcaccaac ggccagatgg ttgatgacat 720  
 cccgcgggag cagcagcgc tcgtgccgt ctacatgagc ggcggcgagg gcgggggcaa 780  
 gaggatccac ccgctccctt tcgcagatcc caacctcca gtgcaaccga gatccatgga 840  
 ccgctccaag gatctggccg cctacggata tggcagcgtg gcctggaagg agagaatgga 900  
 gggctggaag cagaagcagg agcgctgca gcatgtcagg agcgagggtg gcggtgattg 960  
 ggatggcgac gatgcagatc tgccactaat ggatgaagct aggcagccat tgtccagaaa 1020  
 agtccctata tcatcaagcc gaattaatcc ctacaggatg attatcgta tccggttgg 1080  
 ggttttgggt ttcttcttc actaccgagt gatgcatccg gcgaaagatg catttgcat 1140  
 gtggctcata tctgtaatct gtgaaatctg gtttgcgatg tcctggattc ttgatcagtt 1200  
 cccaaagtgg cttccaatcg agagagagc ttacctggac cgtttgtcac taaggtttga 1260  
 caaggaaggt caaccctctc agcttgctcc aatcgacttc tttgtcagta cggttgatcc 1320  
 cacaaaggaa cctcccttgg tcacagcgaa cactgtcctt tccatccttt ctgtggatta 1380  
 tccggttgag aaggtctcct gctatgtttc tgatgatggt gctgcaatgc ttacgtttga 1440  
 agcattgtct gaaacatctg aatttgcaaa gaaatgggtt ctttcagca aaaagttaa 1500  
 tatcgagcct cgtgctcctg agtggtactt ccaacagaag atagactacc tgaagacaa 1560  
 ggttgctgct tcatttgta gggagaggag ggcgatgaag agagaatacg aggaattcaa 1620  
 ggtaaggatc aatgccttgg ttgcaaaagc ccaaagggtt cctgaggaag gatggacaat 1680  
 gcaagatgga agcccctggc ctggaaacaa cgtacgcgat catcctggaa tgattcaggt 1740  
 attccttggc caaagtggcg gtcgtgatgt ggaaggaaat gagttgcctc gcctggttta 1800  
 tgtctcgaga gaaaagaggc caggttataa ccatcacaag aaggctggtg ccatgaatgc 1860  
 actggtccgt gtctctgctg tcttatcaaa tgctgcatac ctattgaact tggactgtga 1920  
 tcactacatc aacaatagca aggccataaa agaggctatg tgtttcatga tggatccttt 1980  
 ggtggggaag aaagtgtgct atgtacagtt ccctcagagg tttgatggta ttgacaaaaa 2040  
 tgatcgatac gctaacagga acgttgtctt tttgacatc aacatgaaag gtttgacagg 2100  
 tattcaagga cccatttatg tgggtactgg atgtgttttc agacggcagg cactgtatgg 2160  
 ttatgatgct cctaaaacga agaagccacc atcaagaact tgcaactgct ggcccaagt 2220  
 gtgcctctct tgcgtgctga gcaggaacaa gaataaaaag aagactacaa aaccaagac 2280  
 ggagaagaag aaaagattat tttcaagaa agcagaaaac ccatctcctg catatgcttt 2340

```

gggtgaaatt gatgaagggtg ctccaggtgc tgatatcgag aaggccggaa tcgtaaatca 2400
acagaaacta gagaagaaat ttgggcagtc ttctgttttt gtcgcatcaa cacttcttga 2460
gaacggaggg accctgaaga gcgcaagtcc agcttctctt ctgaaggaag ctatacatgt 2520
tatcagctgc ggctacgaag acaagaccga ctggggaaaa gagattggct ggattttacgg 2580
atcgatcaca gaggatatct tgactggatt taagatgcac tgccatggct ggcggtctat 2640
ttactgcatc ccgaagcggc ctgcattcaa aggttctgcg cctctgaacc ttccgaccg 2700
tcttcaccag gtccttcgct gggcccttgg gtccgtcgaa attttcttca gcaagcactg 2760
cccactttgg tacggatacg gcggcgggct aaaattcctg gaaaggtttt cttatatcaa 2820
ctccatcgtt tatccctgga cgtccattcc tctcctggct tactgtacct tgcctgccat 2880
ctgcctgtc acggggaagt ttatcacacc agagcttacc aatgtcgcca gtatctggtt 2940
catggcactt ttcattctga tctccgtgac cggcaccctg gaaatgaggt ggagtggcgt 3000
ggccatcgac gactggtgga ggaacgagca gttctgggtc atcggaggcg tttcggcgca 3060
tctgttcgag gtgttcagg gcctgctgaa ggtgttcgcc ggcacgaca cgagcttcac 3120
cgtgacgtcg aaggccggg acgacgagga ctgtcggag ctgtacacgt tcaagtggac 3180
caccctgtcg atacccccga ccacgctcct cctgctgaac ttcacgggg tgggtggccg 3240
gatctcgaac gcgatcaaca acgggtacga gtcgtggggc cccctgttcg ggaagctctt 3300
cttcgccttc tgggtgatcg tccacctgta cccgttcctc aagggtctgg tggggaggca 3360
gaacaggacg ccgacgatcg tcatcgtctg gtccatcctg ctggcctcga tcttctcgt 3420
cctgtgggtc cgcgtcgacc cgttcctcgc caagagcaac ggcccgcctc tggaggagt 3480
tggcctggac tgcaactgaa gtggggggcc cctgtcactc gaagttctgt cacgggcgaa 3540
ttacgcctga tttttgttg ttgttgtgt tggaattctt tgctgtagat agaaaccaca 3600
tgtccacggc atctctgctg tgtccattgg agcaggagag aggtgcctgc tgctgtttgt 3660
tgagtaaatt aaaagtttta aagttataca gtgatgcaca ttccagtgcc cagtgtattc 3720
cctttttaca gtctgtatat tagcgacaaa ggacatattg gttaggagtt tgattctttt 3780
gtaaaa 3786

```

<210> 8  
 <211> 1165  
 <212> PRT  
 <213> Zea mays

<400> 8  
 His Ser Ser Tyr Thr Lys Ser Arg Ser Ser Leu Ala Gln Pro Arg Ala  
 1 5 10 15  
 Ala Pro Arg Gln Ala Gln Pro Pro Pro Ala Thr Ala Ala Cys Ala Cys  
 20 25 30  
 Glu Arg Ser Pro Arg Pro Gly Asp Gln Arg Arg Gly Gly Leu Arg Ala  
 35 40 45  
 Phe Arg Cys Ala Ala Ala Ala Gly Phe Val Arg Glu Arg Asp Pro Ala  
 50 55 60  
 Gly Arg Gly Gly Gly Pro Glu Met Glu Ala Ser Ala Gly Leu Val Ala  
 65 70 75 80  
 Gly Ser His Asn Arg Asn Glu Leu Val Val Ile Arg Arg Asp Arg Glu  
 85 90 95  
 Ser Gly Ala Ala Gly Gly Gly Ala Ala Arg Arg Ala Glu Ala Pro Cys  
 100 105 110  
 Gln Ile Cys Gly Asp Glu Val Gly Val Gly Phe Asp Gly Glu Pro Phe  
 115 120 125  
 Val Ala Cys Asn Glu Cys Ala Phe Pro Val Cys Arg Ala Cys Tyr Glu  
 130 135 140

Tyr Glu Arg Arg Glu Gly Ser Gln Ala Cys Pro Gln Cys Arg Thr Arg  
 145 150 155 160  
 Tyr Lys Arg Leu Lys Gly Cys Pro Arg Val Ala Gly Asp Glu Glu Glu  
 165 170 175  
 Asp Gly Val Asp Asp Leu Glu Gly Glu Phe Gly Leu Gln Asp Gly Ala  
 180 185 190  
 Ala His Glu Asp Asp Pro Gln Tyr Val Ala Glu Ser Met Leu Arg Ala  
 195 200 205  
 Gln Met Ser Tyr Gly Arg Gly Gly Asp Ala His Pro Gly Phe Ser Pro  
 210 215 220  
 Val Pro Asn Val Pro Leu Leu Thr Asn Gly Gln Met Val Asp Asp Ile  
 225 230 235 240  
 Pro Pro Glu Gln His Ala Leu Val Pro Ser Tyr Met Ser Gly Gly Gly  
 245 250 255  
 Gly Gly Gly Lys Arg Ile His Pro Leu Pro Phe Ala Asp Pro Asn Leu  
 260 265 270  
 Pro Val Gln Pro Arg Ser Met Asp Pro Ser Lys Asp Leu Ala Ala Tyr  
 275 280 285  
 Gly Tyr Gly Ser Val Ala Trp Lys Glu Arg Met Glu Gly Trp Lys Gln  
 290 295 300  
 Lys Gln Glu Arg Leu Gln His Val Arg Ser Glu Gly Gly Gly Asp Trp  
 305 310 315 320  
 Asp Gly Asp Asp Ala Asp Leu Pro Leu Met Asp Glu Ala Arg Gln Pro  
 325 330 335  
 Leu Ser Arg Lys Val Pro Ile Ser Ser Ser Arg Ile Asn Pro Tyr Arg  
 340 345 350  
 Met Ile Ile Val Ile Arg Leu Val Val Leu Gly Phe Phe Phe His Tyr  
 355 360 365  
 Arg Val Met His Pro Ala Lys Asp Ala Phe Ala Leu Trp Leu Ile Ser  
 370 375 380  
 Val Ile Cys Glu Ile Trp Phe Ala Met Ser Trp Ile Leu Asp Gln Phe  
 385 390 395 400  
 Pro Lys Trp Leu Pro Ile Glu Arg Glu Thr Tyr Leu Asp Arg Leu Ser  
 405 410 415  
 Leu Arg Phe Asp Lys Glu Gly Gln Pro Ser Gln Leu Ala Pro Ile Asp  
 420 425 430  
 Phe Phe Val Ser Thr Val Asp Pro Thr Lys Glu Pro Pro Leu Val Thr  
 435 440 445  
 Ala Asn Thr Val Leu Ser Ile Leu Ser Val Asp Tyr Pro Val Glu Lys  
 450 455 460

Val Ser Cys Tyr Val Ser Asp Asp Gly Ala Ala Met Leu Thr Phe Glu  
 465 470 475 480  
 Ala Leu Ser Glu Thr Ser Glu Phe Ala Lys Lys Trp Val Pro Phe Ser  
 485 490 495  
 Lys Lys Phe Asn Ile Glu Pro Arg Ala Pro Glu Trp Tyr Phe Gln Gln  
 500 505 510  
 Lys Ile Asp Tyr Leu Lys Asp Lys Val Ala Ala Ser Phe Val Arg Glu  
 515 520 525  
 Arg Arg Ala Met Lys Arg Glu Tyr Glu Glu Phe Lys Val Arg Ile Asn  
 530 535 540  
 Ala Leu Val Ala Lys Ala Gln Lys Val Pro Glu Glu Gly Trp Thr Met  
 545 550 555 560  
 Gln Asp Gly Ser Pro Trp Pro Gly Asn Asn Val Arg Asp His Pro Gly  
 565 570 575  
 Met Ile Gln Val Phe Leu Gly Gln Ser Gly Gly Arg Asp Val Glu Gly  
 580 585 590  
 Asn Glu Leu Pro Arg Leu Val Tyr Val Ser Arg Glu Lys Arg Pro Gly  
 595 600 605  
 Tyr Asn His His Lys Lys Ala Gly Ala Met Asn Ala Leu Val Arg Val  
 610 615 620  
 Ser Ala Val Leu Ser Asn Ala Ala Tyr Leu Leu Asn Leu Asp Cys Asp  
 625 630 635 640  
 His Tyr Ile Asn Asn Ser Lys Ala Ile Lys Glu Ala Met Cys Phe Met  
 645 650 655  
 Met Asp Pro Leu Val Gly Lys Lys Val Cys Tyr Val Gln Phe Pro Gln  
 660 665 670  
 Arg Phe Asp Gly Ile Asp Lys Asn Asp Arg Tyr Ala Asn Arg Asn Val  
 675 680 685  
 Val Phe Phe Asp Ile Asn Met Lys Gly Leu Asp Gly Ile Gln Gly Pro  
 690 695 700  
 Ile Tyr Val Gly Thr Gly Cys Val Phe Arg Arg Gln Ala Leu Tyr Gly  
 705 710 715 720  
 Tyr Asp Ala Pro Lys Thr Lys Lys Pro Pro Ser Arg Thr Cys Asn Cys  
 725 730 735  
 Trp Pro Lys Trp Cys Leu Ser Cys Cys Cys Ser Arg Asn Lys Asn Lys  
 740 745 750  
 Lys Lys Thr Thr Lys Pro Lys Thr Glu Lys Lys Lys Arg Leu Phe Phe  
 755 760 765  
 Lys Lys Ala Glu Asn Pro Ser Pro Ala Tyr Ala Leu Gly Glu Ile Asp  
 770 775 780

Glu Gly Ala Pro Gly Ala Asp Ile Glu Lys Ala Gly Ile Val Asn Gln  
 785 790 795 800  
 Gln Lys Leu Glu Lys Lys Phe Gly Gln Ser Ser Val Phe Val Ala Ser  
 805 810 815  
 Thr Leu Leu Glu Asn Gly Gly Thr Leu Lys Ser Ala Ser Pro Ala Ser  
 820 825 830  
 Leu Leu Lys Glu Ala Ile His Val Ile Ser Cys Gly Tyr Glu Asp Lys  
 835 840 845  
 Thr Asp Trp Gly Lys Glu Ile Gly Trp Ile Tyr Gly Ser Ile Thr Glu  
 850 855 860  
 Asp Ile Leu Thr Gly Phe Lys Met His Cys His Gly Trp Arg Ser Ile  
 865 870 875 880  
 Tyr Cys Ile Pro Lys Arg Pro Ala Phe Lys Gly Ser Ala Pro Leu Asn  
 885 890 895  
 Leu Ser Asp Arg Leu His Gln Val Leu Arg Trp Ala Leu Gly Ser Val  
 900 905 910  
 Glu Ile Phe Phe Ser Lys His Cys Pro Leu Trp Tyr Gly Tyr Gly Gly  
 915 920 925  
 Gly Leu Lys Phe Leu Glu Arg Phe Ser Tyr Ile Asn Ser Ile Val Tyr  
 930 935 940  
 Pro Trp Thr Ser Ile Pro Leu Leu Ala Tyr Cys Thr Leu Pro Ala Ile  
 945 950 955 960  
 Cys Leu Leu Thr Gly Lys Phe Ile Thr Pro Glu Leu Thr Asn Val Ala  
 965 970 975  
 Ser Ile Trp Phe Met Ala Leu Phe Ile Cys Ile Ser Val Thr Gly Ile  
 980 985 990  
 Leu Glu Met Arg Trp Ser Gly Val Ala Ile Asp Asp Trp Trp Arg Asn  
 995 1000 1005  
 Glu Gln Phe Trp Val Ile Gly Gly Val Ser Ala His Leu Phe Ala Val  
 1010 1015 1020  
 Phe Gln Gly Leu Leu Lys Val Phe Ala Gly Ile Asp Thr Ser Phe Thr  
 1025 1030 1035 1040  
 Val Thr Ser Lys Ala Gly Asp Asp Glu Glu Phe Ser Glu Leu Tyr Thr  
 1045 1050 1055  
 Phe Lys Trp Thr Thr Leu Leu Ile Pro Pro Thr Thr Leu Leu Leu Leu  
 1060 1065 1070  
 Asn Phe Ile Gly Val Val Ala Gly Ile Ser Asn Ala Ile Asn Asn Gly  
 1075 1080 1085  
 Tyr Glu Ser Trp Gly Pro Leu Phe Gly Lys Leu Phe Phe Ala Phe Trp  
 1090 1095 1100

Val Ile Val His Leu Tyr Pro Phe Leu Lys Gly Leu Val Gly Arg Gln  
 1105 1110 1115 1120

Asn Arg Thr Pro Thr Ile Val Ile Val Trp Ser Ile Leu Leu Ala Ser  
 1125 1130 1135

Ile Phe Ser Leu Leu Trp Val Arg Val Asp Pro Phe Leu Ala Lys Ser  
 1140 1145 1150

Asn Gly Pro Leu Leu Glu Glu Cys Gly Leu Asp Cys Asn  
 1155 1160 1165

<210> 9  
 <211> 3936  
 <212> DNA  
 <213> Zea mays

<400> 9

cttctccctc	gtcgggtgcgg	cgtgggcgcgg	ctcgggcgttc	ggtgagaaac	cactcggggg	60
atgaggatct	gctgctagag	tgagaggagc	tacggtcagt	atcctctgcc	ttcgtcggcg	120
gcggaagtgg	aggggaggaa	gcgatggagg	cgagcgccgg	gctgggtggc	ggctcccaca	180
accgcaacga	gctcgtcgtc	atccgcccgc	acggcgatcc	cgggccgaag	ccgccgcggg	240
agcagaacgg	gcaggtgtgc	cagatttgcg	gcgacgacgt	cggccttgcc	cccggcgggg	300
accccttcgt	ggcgtgcaac	gagtgcgcc	tccccgtctg	ccgggactgc	tacgaatacg	360
agcgccggga	gggcacgcag	aactgcccc	agtgaagac	tcgatacaag	cgcctcaagg	420
gctgccaacg	tgtgaccgg	gacgaggagg	aggacggcgt	cgatgacctg	gacaacgagt	480
tcaactggga	acgaactgac	tcgcagtcgt	tgcccgagtc	catgctctac	ggccacatga	540
gctacggccg	tggagggtgac	cctaattggc	cgccacaagc	ttccagctc	aaccccaatg	600
ttccactcct	caccaacggg	caaattggtg	atgacatccc	accggagcag	cacgcgctgg	660
tgccttcttt	catgggtggt	gggggaaaga	ggatacatcc	ccttccttat	gcggatccca	720
gcttacctgt	gcaacccagg	tctatggacc	catccaagga	tcttgctgca	tatgggtatg	780
gtagtggtgc	ttggaaggaa	cggatggaga	attggaagca	gagacaagag	aggatgcacc	840
agacggggaa	tgatggtggt	ggtgatgatg	gtgacgatgc	tgatctacca	ctaattggatg	900
aagcaagaca	acaactgtcc	aggaataattc	cacttccatc	aagccagatt	aatccatata	960
ggatgattat	cattattcgg	cttgtggttt	tggggttctt	cttccactac	cgagtgatgc	1020
atccggtgaa	tgatgcattt	gctttgtggc	tcatactctg	tatctgtgaa	atctggtttg	1080
ccatgtcttg	gattcttgat	caattcccaa	agtggttccc	tattgagaga	gagacttacc	1140
tagaccggct	gtcactgagg	ttcgacaagg	aaggccagcc	atctcaactt	gctccaattg	1200
atctctttgt	cagtacggtt	gatcccttaa	aggaacctcc	tttggtcaca	acaaatactg	1260
ttctatctat	cctttcggtg	gattatcctg	ttgataaggt	ttcttgctat	gtttctgatg	1320
atggtgctgc	aattgctaag	tttgaagcat	tatctgaaac	atctgaattt	gcaaagaaat	1380
gggttccttt	ctgcaaacgg	tacaatattg	aacctcgcgc	tccagagtgg	tacttccaa	1440
agaagataga	ctacttgaaa	gacaagggtg	cagcaaaactt	tgtaggggag	aggagagcaa	1500
tgaagagaga	gtatgaggaa	ttcaagggtg	gaatcaatgc	cttagttgcc	aaagcccaga	1560
aagttcctga	agaaggatgg	acaatgcaag	atggaacccc	ctggcctgga	aacaatgttc	1620
gtgatcatcc	tggaatgatt	caggctctcc	ttggccaaag	cggaggcctt	gactgtgagg	1680
gaaatgaact	gccacgattg	gtttatgttt	ctagagagaa	acgaccaggc	tataaccatc	1740
ataagaaaagc	tggtgctatg	aatgcattgg	tccgagtcct	tgctgtacta	acaaatgctc	1800
catatttggt	aaacttggat	tgtgatcact	acatcaacaa	cagcaaggct	ataaagggaag	1860
caatgtgttt	tatgatggac	cctttactag	gaaagaagggt	ttgctatgta	cagttccctc	1920
aaagatttga	tgggattgat	cgccatgacc	gatatgctaa	ccggaatgtt	gtcttttttg	1980
atatcaacat	gaaaggtttg	gatggtattc	agggctcaat	ttatgttggg	actggatgtg	2040
tatttagaag	gcaggcatta	tatggttatg	atgcccccaa	aacaaagaag	ccaccatcaa	2100
ggacttgcaa	ctgctggccc	aagtgggtgt	tttgctgttg	ctgctttggc	aataggaagc	2160
aaaagaagac	taccaaaccc	aaaacagaga	agaaaaagtt	attatttttc	aagaaagaag	2220
agaaccaatc	ccctgcatat	gctcttggtg	aaattgacga	agctgtccca	ggagctgaga	2280
atgaaaaggc	cggatttgta	aatcaacaaa	aattagaaaa	gaaatttggc	caatcttctg	2340
tttttggtac	atccacactt	ctcgagaatg	gtggaacctt	gaagagtgca	agtcctgctt	2400
ctcttttgaa	agaagctata	catgtcatta	gttgtggtta	tgaagacaag	acagactggg	2460
gaaaagagat	tggctggatc	tatggatcag	ttacagaaga	tattctaact	ggtttcaaga	2520

```

tgcattgtca tgggtggcgg tcaatttact gcatacctaa acggggttga ttcaaagggt 2580
ctgcacctct gaatctttca gatcgtcttc accagggtgct tcgggtgggct cttgggtcta 2640
ttgagatctt cttcagcaat cattgccctc tttggtatgg gtatggtggc ggtctgaaat 2700
ttttggaaag attttcctac atcaactcca tcgtgtatcc ttggacatct attccctct 2760
tggcttactg tacattgcct gccatctgtt tattgacagg gaaattttat actccagagc 2820
tgaataatgt tgccagcctg tggttcatgt cactttttat ctgcattttt gctacgagca 2880
tcctagaaat gagatggagt ggtgttgaa ttgatgactg gtggaggaat gagcagttct 2940
gggtcattgg aggtgtgtcc tcacacctct ttgctgtgtt ccagggactt ctcaagggtca 3000
tagctgggtg tgatacaagc ttcaccgtga catcaaaggg tggagatgat gaggagttct 3060
cagagctata tacattcaaa tggactacct tattgatacc tcctaccacc ttgcttctat 3120
tgaacttcat tgggtgtggtc gctggcggtt caaatgcgat caataacgga tatgagtcac 3180
ggggcccccct ctttgggaag ctattctttg cattttgggt gattgtccat ctttatccct 3240
ttctcaaagg tttgggttga aggcacaaac ggacaccaaac gattgtcatc gtctggtcca 3300
ttctgctggc ttcaatcttc tcgtcctttt ggggttcggat tgatcctttc cttgcgaagg 3360
atgatggtcc gcttcttgag gagtgtggtt tggattgcaa ctaggatgtc agtgcacag 3420
ctccccaat ctgcatatgc ttgaagtata tttctggtg tttgtcccca tattcagtg 3480
ctgtagataa gagacatgaa atgtcccaag tttcttttga tccatggtga acctacttaa 3540
tatctgagag atatactggg ggaaaatgga ggctgcgga atccttggtc agtggggccg 3600
tggaaatacag catatgcaag tgtttgattg tgcagcattc tttattactt ggtcgcaata 3660
tagatgggct gagccgaaca gcaaggattt ttgattctgc actgctcccg tgtacaaact 3720
tggttctcaa taaggcaggc aggaatgcat ctgccagtgg aacagagcaa cctgcacatt 3780
atztatgtat gcctgttcat tggagggtt gttcattaca tgttcgtcta tactagaaaa 3840
aacagaatat tagcattaat ctatagttaa ttaaagtatg taaatgcgcc tgttttttgt 3900
tgtgtactgt aatcatctga gttggtttt tgaaaa 3936

```

<210> 10  
 <211> 1086  
 <212> PRT  
 <213> Zea mays

<400> 10  
 Met Glu Ala Ser Ala Gly Leu Val Ala Gly Ser His Asn Arg Asn Glu  
 1 5 10 15  
 Leu Val Val Ile Arg Arg Asp Gly Asp Pro Gly Pro Lys Pro Pro Arg  
 20 25 30  
 Glu Gln Asn Gly Gln Val Cys Gln Ile Cys Gly Asp Asp Val Gly Leu  
 35 40 45  
 Ala Pro Gly Gly Asp Pro Phe Val Ala Cys Asn Glu Cys Ala Phe Pro  
 50 55 60  
 Val Cys Arg Asp Cys Tyr Glu Tyr Glu Arg Arg Glu Gly Thr Gln Asn  
 65 70 75 80  
 Cys Pro Gln Cys Lys Thr Arg Tyr Lys Arg Leu Lys Gly Cys Gln Arg  
 85 90 95  
 Val Thr Gly Asp Glu Glu Glu Asp Gly Val Asp Asp Leu Asp Asn Glu  
 100 105 110  
 Phe Asn Trp Asp Gly His Asp Ser Gln Ser Val Ala Glu Ser Met Leu  
 115 120 125  
 Tyr Gly His Met Ser Tyr Gly Arg Gly Gly Asp Pro Asn Gly Ala Pro  
 130 135 140  
 Gln Ala Phe Gln Leu Asn Pro Asn Val Pro Leu Leu Thr Asn Gly Gln  
 145 150 155 160



Met Val Asp Asp Ile Pro Pro Glu Gln His Ala Leu Val Pro Ser Phe  
 165 170 175  
 Met Gly Gly Gly Gly Lys Arg Ile His Pro Leu Pro Tyr Ala Asp Pro  
 180 185 190  
 Ser Leu Pro Val Gln Pro Arg Ser Met Asp Pro Ser Lys Asp Leu Ala  
 195 200 205  
 Ala Tyr Gly Tyr Gly Ser Val Ala Trp Lys Glu Arg Met Glu Asn Trp  
 210 215 220  
 Lys Gln Arg Gln Glu Arg Met His Gln Thr Gly Asn Asp Gly Gly Gly  
 225 230 235 240  
 Asp Asp Gly Asp Asp Ala Asp Leu Pro Leu Met Asp Glu Ala Arg Gln  
 245 250 255  
 Gln Leu Ser Arg Lys Ile Pro Leu Pro Ser Ser Gln Ile Asn Pro Tyr  
 260 265 270  
 Arg Met Ile Ile Ile Ile Arg Leu Val Val Leu Gly Phe Phe Phe His  
 275 280 285  
 Tyr Arg Val Met His Pro Val Asn Asp Ala Phe Ala Leu Trp Leu Ile  
 290 295 300  
 Ser Val Ile Cys Glu Ile Trp Phe Ala Met Ser Trp Ile Leu Asp Gln  
 305 310 315 320  
 Phe Pro Lys Trp Phe Pro Ile Glu Arg Glu Thr Tyr Leu Asp Arg Leu  
 325 330 335  
 Ser Leu Arg Phe Asp Lys Glu Gly Gln Pro Ser Gln Leu Ala Pro Ile  
 340 345 350  
 Asp Phe Phe Val Ser Thr Val Asp Pro Leu Lys Glu Pro Pro Leu Val  
 355 360 365  
 Thr Thr Asn Thr Val Leu Ser Ile Leu Ser Val Asp Tyr Pro Val Asp  
 370 375 380  
 Lys Val Ser Cys Tyr Val Ser Asp Asp Gly Ala Ala Met Leu Thr Phe  
 385 390 395 400  
 Glu Ala Leu Ser Glu Thr Ser Glu Phe Ala Lys Lys Trp Val Pro Phe  
 405 410 415  
 Cys Lys Arg Tyr Asn Ile Glu Pro Arg Ala Pro Glu Trp Tyr Phe Gln  
 420 425 430  
 Gln Lys Ile Asp Tyr Leu Lys Asp Lys Val Ala Ala Asn Phe Val Arg  
 435 440 445  
 Glu Arg Arg Ala Met Lys Arg Glu Tyr Glu Glu Phe Lys Val Arg Ile  
 450 455 460  
 Asn Ala Leu Val Ala Lys Ala Gln Lys Val Pro Glu Glu Gly Trp Thr  
 465 470 475 480

Met Gln Asp Gly Thr Pro Trp Pro Gly Asn Asn Val Arg Asp His Pro  
 485 490 495  
 Gly Met Ile Gln Val Phe Leu Gly Gln Ser Gly Gly Leu Asp Cys Glu  
 500 505 510  
 Gly Asn Glu Leu Pro Arg Leu Val Tyr Val Ser Arg Glu Lys Arg Pro  
 515 520 525  
 Gly Tyr Asn His His Lys Lys Ala Gly Ala Met Asn Ala Leu Val Arg  
 530 535 540  
 Val Ser Ala Val Leu Thr Asn Ala Pro Tyr Leu Leu Asn Leu Asp Cys  
 545 550 555 560  
 Asp His Tyr Ile Asn Asn Ser Lys Ala Ile Lys Glu Ala Met Cys Phe  
 565 570 575  
 Met Met Asp Pro Leu Leu Gly Lys Lys Val Cys Tyr Val Gln Phe Pro  
 580 585 590  
 Gln Arg Phe Asp Gly Ile Asp Arg His Asp Arg Tyr Ala Asn Arg Asn  
 595 600 605  
 Val Val Phe Phe Asp Ile Asn Met Lys Gly Leu Asp Gly Ile Gln Gly  
 610 615 620  
 Pro Ile Tyr Val Gly Thr Gly Cys Val Phe Arg Arg Gln Ala Leu Tyr  
 625 630 635 640  
 Gly Tyr Asp Ala Pro Lys Thr Lys Lys Pro Pro Ser Arg Thr Cys Asn  
 645 650 655  
 Cys Trp Pro Lys Trp Cys Phe Cys Cys Cys Cys Phe Gly Asn Arg Lys  
 660 665 670  
 Gln Lys Lys Thr Thr Lys Pro Lys Thr Glu Lys Lys Lys Leu Leu Phe  
 675 680 685  
 Phe Lys Lys Glu Glu Asn Gln Ser Pro Ala Tyr Ala Leu Gly Glu Ile  
 690 695 700  
 Asp Glu Ala Ala Pro Gly Ala Glu Asn Glu Lys Ala Gly Ile Val Asn  
 705 710 715 720  
 Gln Gln Lys Leu Glu Lys Lys Phe Gly Gln Ser Ser Val Phe Val Thr  
 725 730 735  
 Ser Thr Leu Leu Glu Asn Gly Gly Thr Leu Lys Ser Ala Ser Pro Ala  
 740 745 750  
 Ser Leu Leu Lys Glu Ala Ile His Val Ile Ser Cys Gly Tyr Glu Asp  
 755 760 765  
 Lys Thr Asp Trp Gly Lys Glu Ile Gly Trp Ile Tyr Gly Ser Val Thr  
 770 775 780  
 Glu Asp Ile Leu Thr Gly Phe Lys Met His Cys His Gly Trp Arg Ser  
 785 790 795 800

Ile Tyr Cys Ile Pro Lys Arg Val Ala Phe Lys Gly Ser Ala Pro Leu  
 805 810 815  
 Asn Leu Ser Asp Arg Leu His Gln Val Leu Arg Trp Ala Leu Gly Ser  
 820 825 830  
 Ile Glu Ile Phe Phe Ser Asn His Cys Pro Leu Trp Tyr Gly Tyr Gly  
 835 840 845  
 Gly Gly Leu Lys Phe Leu Glu Arg Phe Ser Tyr Ile Asn Ser Ile Val  
 850 855 860  
 Tyr Pro Trp Thr Ser Ile Pro Leu Leu Ala Tyr Cys Thr Leu Pro Ala  
 865 870 875 880  
 Ile Cys Leu Leu Thr Gly Lys Phe Ile Thr Pro Glu Leu Asn Asn Val  
 885 890 895  
 Ala Ser Leu Trp Phe Met Ser Leu Phe Ile Cys Ile Phe Ala Thr Ser  
 900 905 910  
 Ile Leu Glu Met Arg Trp Ser Gly Val Gly Ile Asp Asp Trp Trp Arg  
 915 920 925  
 Asn Glu Gln Phe Trp Val Ile Gly Gly Val Ser Ser His Leu Phe Ala  
 930 935 940  
 Val Phe Gln Gly Leu Leu Lys Val Ile Ala Gly Val Asp Thr Ser Phe  
 945 950 955 960  
 Thr Val Thr Ser Lys Gly Gly Asp Asp Glu Glu Phe Ser Glu Leu Tyr  
 965 970 975  
 Thr Phe Lys Trp Thr Thr Leu Leu Ile Pro Pro Thr Thr Leu Leu Leu  
 980 985 990  
 Leu Asn Phe Ile Gly Val Val Ala Gly Val Ser Asn Ala Ile Asn Asn  
 995 1000 1005  
 Gly Tyr Glu Ser Trp Gly Pro Leu Phe Gly Lys Leu Phe Phe Ala Phe  
 1010 1015 1020  
 Trp Val Ile Val His Leu Tyr Pro Phe Leu Lys Gly Leu Val Gly Arg  
 1025 1030 1035 1040  
 Gln Asn Arg Thr Pro Thr Ile Val Ile Val Trp Ser Ile Leu Leu Ala  
 1045 1050 1055  
 Ser Ile Phe Ser Leu Leu Trp Val Arg Ile Asp Pro Phe Leu Ala Lys  
 1060 1065 1070  
 Asp Asp Gly Pro Leu Leu Glu Glu Cys Gly Leu Asp Cys Asn  
 1075 1080 1085

<210> 11  
 <211> 1138  
 <212> DNA  
 <213> *Oryza sativa*

<400> 11  
 cgctgctccc ggcgatggac gtgttcgtca ccaccgccga ccccgacaag gagccgccgc 60  
 tcgccacggc gaacaccgtg ctgtccatat atcctcgccg cgggctaccc cgccggcaag 120  
 gtgacgtgct atatttccga cgacgcaggc gcggagggtga cacgtaacgc ggtcgtggag 180  
 gcggcccggg tcgcggcgct ttgggtgtcg ttctgccgga agcacggcgt cgagccgagg 240  
 aacctggagg cgtacttcaa cgccggcgag ggtggtggtg gcaaggcgaa ggtggtggcg 300  
 agggggagct acagggggat ggctgggccg gagctggtgc gcgacaggag acgggtgcgc 360  
 cgcgagtacg aggagatgcg gctgcggatc gacgcgctgc aggccgccga tgcgcgccgc 420  
 cggcgccgcg gcgcggccga tgaccacgcc ggagttgtgc aggtactgat cgattttgct 480  
 gggagcgtgc cacagctcgg cgttgccaac gggagcaagc tcatcgacgt cgctctgtc 540  
 gacgtgtgcc tcccggcgct tgtgtacgtg tgccgcgaga agcgccgcgg ccacgcgcac 600  
 caccggaagg cgggcgccat gaacgcgcc ttcctcctcg acctcgactg cgactactac 660  
 gtcaacaact cgcaggccct cgccggccgc atctgcttca tgatcgaacg cggcggcgcc 720  
 ggagccgccg aagacgccgg cgcggtcgcg ttcgtccagt tcccgcagcg ggtcgacggc 780  
 gtcgatcccg gcgaccgcta cgccaaccac aaccgcgtcc tcttcgactg caccgagctc 840  
 ggcctcgacg gcctccaggg ccccatctac gtcggcaccg gctgcttgtt ccgccgtgtc 900  
 gcgctctaca gcgtcgacct gccgcgctgg agaccgcggc gttcattggg ctgtcgccta 960  
 ctcggagaag acgagcggct atggtccagg atgaaacaaa tggtaatat aagtgggtcca 1020  
 aggtgaaaaa ctcagctaaa acctgaccga agctgtaaca tgggtaaaaa tatatggccc 1080  
 aaaatgaaat ttactttttt ttttttacca aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1138

<210> 12  
 <211> 341  
 <212> PRT  
 <213> *Oryza sativa*

<400> 12  
 Arg Cys Ser Arg Arg Trp Thr Cys Ser Ser Pro Pro Pro Thr Pro Thr  
 1 5 10 15  
 Arg Ser Arg Arg Ser Pro Arg Arg Thr Pro Cys Cys Pro Tyr Ile Leu  
 20 25 30  
 Ala Ala Gly Tyr Pro Ala Gly Lys Val Thr Cys Tyr Ile Ser Asp Asp  
 35 40 45  
 Ala Gly Ala Glu Val Thr Arg Asn Ala Val Val Glu Ala Ala Arg Phe  
 50 55 60  
 Ala Ala Leu Trp Val Ser Phe Cys Arg Lys His Gly Val Glu Pro Arg  
 65 70 75 80  
 Asn Leu Glu Ala Tyr Phe Asn Ala Gly Glu Gly Gly Gly Gly Lys Ala  
 85 90 95  
 Lys Val Val Ala Arg Gly Ser Tyr Arg Gly Met Ala Trp Pro Glu Leu  
 100 105 110  
 Val Arg Asp Arg Arg Arg Val Arg Arg Glu Tyr Glu Glu Met Arg Leu  
 115 120 125  
 Arg Ile Asp Ala Leu Gln Ala Ala Asp Ala Arg Arg Arg Arg Gly  
 130 135 140  
 Ala Ala Asp Asp His Ala Gly Val Val Gln Val Leu Ile Asp Phe Ala  
 145 150 155 160  
 Gly Ser Val Pro Gln Leu Gly Val Ala Asn Gly Ser Lys Leu Ile Asp  
 165 170 175

Val Ala Ser Val Asp Val Cys Leu Pro Ala Leu Val Tyr Val Cys Arg  
 180 185 190  
 Glu Lys Arg Arg Gly His Ala His His Arg Lys Ala Gly Ala Met Asn  
 195 200 205  
 Ala Pro Phe Ile Leu Asp Leu Asp Cys Asp Tyr Tyr Val Asn Asn Ser  
 210 215 220  
 Gln Ala Leu Arg Ala Gly Ile Cys Phe Met Ile Glu Arg Gly Gly Gly  
 225 230 235 240  
 Gly Ala Ala Glu Asp Ala Gly Ala Val Ala Phe Val Gln Phe Pro Gln  
 245 250 255  
 Arg Val Asp Gly Val Asp Pro Gly Asp Arg Tyr Ala Asn His Asn Arg  
 260 265 270  
 Val Leu Phe Asp Cys Thr Glu Leu Gly Leu Asp Gly Leu Gln Gly Pro  
 275 280 285  
 Ile Tyr Val Gly Thr Gly Cys Leu Phe Arg Arg Val Ala Leu Tyr Ser  
 290 295 300  
 Val Asp Leu Pro Arg Trp Arg Pro Arg Arg Ser Leu Gly Cys Arg Leu  
 305 310 315 320  
 Leu Gly Glu Asp Glu Arg Leu Trp Ser Arg Met Lys Gln Met Val Ile  
 325 330 335  
 Leu Ser Gly Pro Arg  
 340

<210> 13  
 <211> 3517  
 <212> DNA  
 <213> Glycine max

<400> 13  
 gcacgagcca acaacaacac ccttatgtgg acacattagg tgaggttcaa cagctagcac 60  
 caatcttcct tcataaaaca caaacctttg atcacacaat ctcaccttaa tttgtgttgt 120  
 tgttgtgccca ttcccatatt gtcccatcca ctaagacatg gaagccagcg ctggactggg 180  
 cgctgggtca cataaccgca atgagctagt gtgcattcat ggccatgaag agccgaaggc 240  
 tttgaagaac ttggatgggc aagtgtgtga gatttgtggg gatggcgtgg gactcacggg 300  
 ggatggagac ttgtttgtgg cttgcaatga gtgtgggttt ccagtgtgca ggccttgcta 360  
 tgagtatgaa aggagagaag gaagccacct ttgccacag tgcaaaacca gatacaagcg 420  
 tctcaaaggg agcccccgag tggagggaga tgatgatgaa gaggatgtgg atgatattga 480  
 gcatgaattc aatattgatg agcaaaaagaa caagcatggc caggttgcag aagccatgct 540  
 tcatgggagg atgagctatg gaagagggtcc tgaagatgat gacaattccc agttcccaac 600  
 acctgtcatt gctgggtggc gttctaggcc tgtaagtggg gagttcccaa tatcatctaa 660  
 tgcttatggg gatcagatgt tatcctcttc actgcataaa agagtgcac catatccagt 720  
 gtctgaacct ggaagtgcaa gatgggacga aaaaaaaaga agatggatgg aaagatagaa 780  
 tggatgactg gaaattgcag caaggcaatt tggggcctga accggatgaa gatccagatg 840  
 cagccatggt agatgaagca aggcaaccac tgtcaaggaa agtgccaata gcatccagca 900  
 aaatcaatcc atatagaatg gtgattgtgg cacgtctggt tattcttgct ttcttctc 960  
 gatacagact catgaaccca gtacatgat ccctggggct atggctaacc tctatcatat 1020  
 gtgaaatctg gtttgctttt tcatggattc tggatcagtt tcccaaattg tttccattg 1080  
 atagagagac ctaccttgac cgtctttcca tcaggtatga gcgtgaagg gaaccaaaca 1140  
 tgcttgctcc tgtagatgtt tttgttagta ccgtggatcc catgaaggaa cctcctctgg 1200  
 ttacagcaaa cactgttctt tcaatcttgg ccattggatta cccggttgat aaaatatcat 1260

```

gctacatttc tgatgatgga gcctcaatgt gtacatttga gtccttatca gaaactgcag 1320
agtttgctag aaagtgggta ccgtttttgta agaaattttc catagaacct cgggcacctg 1380
agatgtactt cagcgagaag attgactacc taaaggacaa agtgcaacct acctttgtta 1440
aggagcgctg agctatgaag aggggaatacg aagagtttaa ggtaggatc aatgcacttg 1500
ttgctaaggc ccagaaagt cctcagggag gatggatcat gcaggatggg acaccatggc 1560
caggggaataa cactaaggat catcctggta tgattcaagt gtttcttggt agcagtggag 1620
gtcttgatac tgaaggaaac caacttcctc gccttgttta tgtttccaga gaaaaaggc 1680
ctggttttca acaccacaag aaagctgggt ccatgaatgc tctggttcgg gtatctgctg 1740
ttctcacaaa tgctcctttc atgttgaact tggattgtga tctactatgtc aataacagca 1800
aggctgcccg agaggccatg tgcttcttga tggaccacaa aactgggaag aaggtctgct 1860
atgtccagtt tcctcaaaga tttgatggta ttgatacaca tgatcgttat gccaacagga 1920
acacagtttt ctttgatatt aacatgaagg gtctagatgg tattcaaggt cctgtatatg 1980
tggggactgg atgtgttttc aggaggcaag ctttgatgg ctataatcct cccaagggtc 2040
caaagcgctc aaaaatggta agctgtgatt gttgcccggt ttttggaagc cgcaagaagt 2100
ataaggagaa gaatgatgca aatggagagg ctgcaagcct aaaagggatg gatgatgaca 2160
aagaggtggt gatgtcccaa atgaattttg agaagaaatt tggacaatcc tctatttttg 2220
tgacttctac cttgatggaa gaggggtggg tgccctcttc ttcaagtcca gctgccctgc 2280
ttaaagaagc cattcatgtg attagctgtg gatatgaaga taaaactgaa tggggacttg 2340
agcttggtg gatctatgga tctatcacag aagatattct aacaggtttt aagatgcatt 2400
gccgtgggtg gaggtccatt tattgtatgc caaagagagc tgcattcaag ggtactgctc 2460
ctatcaactt gtcagatcgt ctcaaccagg ttcttcgttg ggcacttggg tccattgaga 2520
ttttctttag tcaccattgc cctctatggg atggcttcaa ggaaaagaag ctaaagtggc 2580
ttgagagatt tgccctatgca aacacaactg tctatccatt cacctccatt cctctagttg 2640
cctactgtat tcttccagca gtttgtttac tctactgacaa attcatcatg ccaccgatta 2700
gcacctttgc tggtttgtac tttgttgctc tcttctcctc aatcattgca actggtattc 2760
ttgagttgaa atggagtgga gtgagcattg aggaatggg gagaaatgag cagttttggg 2820
tcattgggtg tgatcagct cacctctttg ctgttataca aggtctgcta aaggttctgg 2880
ctggaattga caccaatttc actgttacat caaaggcaac agatgatgaa gagtttggag 2940
aattgtacac cttaagtgg actacactct tgattcctcc aaccactatt ttgatcatta 3000
acattgttg tgttgttgct ggaatctcag atgccataaa caatgggtac caatcctggg 3060
gaccactctt tggaaagctc ttcttttctc tctgggtgat tgtccatctc tatccattcc 3120
ttaaagggtt gatgggtcgc caaaatcgca caccacccat tgttgatgatt tggtcagtgc 3180
tattggcctc tattttctcc ttactttggg taagaattga tccatttgtc ctcaagacta 3240
agggacctga taccaagcta tgtggaatca actgctaaaa aagactgctt tccctatagt 3300
attattcttt aaaagatgta tgtagggtac atacattctt ggtttcacaa accaacaag 3360
tggcaatgca caaggatcaa taaggaaaga gtgaaaattt tgtgtatcat aaatgagtgt 3420
tatcattttt gtaaatgttc tcaaggacat ctgttttggg tggaaactgcc caaaaattgc 3480
agttttatct attcactgga aaaaaaaaaa aaaaaaa 3517

```

<210> 14  
 <211> 1039  
 <212> PRT  
 <213> Glycine max

<220>  
 <221> UNSURE  
 <222> (201)

<400> 14  
 Met Glu Ala Ser Ala Gly Leu Val Ala Gly Ser His Asn Arg Asn Glu  
 1 5 10 15  
 Leu Val Val Ile His Gly His Glu Glu Pro Lys Ala Leu Lys Asn Leu  
 20 25 30  
 Asp Gly Gln Val Cys Glu Ile Cys Gly Asp Gly Val Gly Leu Thr Val  
 35 40 45  
 Asp Gly Asp Leu Phe Val Ala Cys Asn Glu Cys Gly Phe Pro Val Cys  
 50 55 60

Arg Pro Cys Tyr Glu Tyr Glu Arg Arg Glu Gly Ser His Leu Cys Pro  
 65 70 75 80  
 Gln Cys Lys Thr Arg Tyr Lys Arg Leu Lys Gly Ser Pro Arg Val Glu  
 85 90 95  
 Gly Asp Asp Asp Glu Glu Asp Val Asp Asp Ile Glu His Glu Phe Asn  
 100 105 110  
 Ile Asp Glu Gln Lys Asn Lys His Gly Gln Val Ala Glu Ala Met Leu  
 115 120 125  
 His Gly Arg Met Ser Tyr Gly Arg Gly Pro Glu Asp Asp Asp Asn Ser  
 130 135 140  
 Gln Phe Pro Thr Pro Val Ile Ala Gly Gly Arg Ser Arg Pro Val Ser  
 145 150 155 160  
 Gly Glu Phe Pro Ile Ser Ser Asn Ala Tyr Gly Asp Gln Met Leu Ser  
 165 170 175  
 Ser Ser Leu His Lys Arg Val His Pro Tyr Pro Val Ser Glu Pro Gly  
 180 185 190  
 Ser Ala Arg Trp Asp Glu Lys Lys Xaa Asp Gly Trp Lys Asp Arg Met  
 195 200 205  
 Asp Asp Trp Lys Leu Gln Gln Gly Asn Leu Gly Pro Glu Pro Asp Glu  
 210 215 220  
 Asp Pro Asp Ala Ala Met Leu Asp Glu Ala Arg Gln Pro Leu Ser Arg  
 225 230 235 240  
 Lys Val Pro Ile Ala Ser Ser Lys Ile Asn Pro Tyr Arg Met Val Ile  
 245 250 255  
 Val Ala Arg Leu Val Ile Leu Ala Phe Phe Leu Arg Tyr Arg Leu Met  
 260 265 270  
 Asn Pro Val His Asp Ala Leu Gly Leu Trp Leu Thr Ser Ile Ile Cys  
 275 280 285  
 Glu Ile Trp Phe Ala Phe Ser Trp Ile Leu Asp Gln Phe Pro Lys Trp  
 290 295 300  
 Phe Pro Ile Asp Arg Glu Thr Tyr Leu Asp Arg Leu Ser Ile Arg Tyr  
 305 310 315 320  
 Glu Arg Glu Gly Glu Pro Asn Met Leu Ala Pro Val Asp Val Phe Val  
 325 330 335  
 Ser Thr Val Asp Pro Met Lys Glu Pro Pro Leu Val Thr Ala Asn Thr  
 340 345 350  
 Val Leu Ser Ile Leu Ala Met Asp Tyr Pro Val Asp Lys Ile Ser Cys  
 355 360 365  
 Tyr Ile Ser Asp Asp Gly Ala Ser Met Cys Thr Phe Glu Ser Leu Ser  
 370 375 380

Glu Thr Ala Glu Phe Ala Arg Lys Trp Val Pro Phe Cys Lys Lys Phe  
 385 390 395 400  
 Ser Ile Glu Pro Arg Ala Pro Glu Met Tyr Phe Ser Glu Lys Ile Asp  
 405 410 415  
 Tyr Leu Lys Asp Lys Val Gln Pro Thr Phe Val Lys Glu Arg Arg Ala  
 420 425 430  
 Met Lys Arg Glu Tyr Glu Glu Phe Lys Val Arg Ile Asn Ala Leu Val  
 435 440 445  
 Ala Lys Ala Gln Lys Val Pro Gln Gly Gly Trp Ile Met Gln Asp Gly  
 450 455 460  
 Thr Pro Trp Pro Gly Asn Asn Thr Lys Asp His Pro Gly Met Ile Gln  
 465 470 475 480  
 Val Phe Leu Gly Ser Ser Gly Gly Leu Asp Thr Glu Gly Asn Gln Leu  
 485 490 495  
 Pro Arg Leu Val Tyr Val Ser Arg Glu Lys Arg Pro Gly Phe Gln His  
 500 505 510  
 His Lys Lys Ala Gly Ala Met Asn Ala Leu Val Arg Val Ser Ala Val  
 515 520 525  
 Leu Thr Asn Ala Pro Phe Met Leu Asn Leu Asp Cys Asp His Tyr Val  
 530 535 540  
 Asn Asn Ser Lys Ala Ala Arg Glu Ala Met Cys Phe Leu Met Asp Pro  
 545 550 555 560  
 Gln Thr Gly Lys Lys Val Cys Tyr Val Gln Phe Pro Gln Arg Phe Asp  
 565 570 575  
 Gly Ile Asp Thr His Asp Arg Tyr Ala Asn Arg Asn Thr Val Phe Phe  
 580 585 590  
 Asp Ile Asn Met Lys Gly Leu Asp Gly Ile Gln Gly Pro Val Tyr Val  
 595 600 605  
 Gly Thr Gly Cys Val Phe Arg Arg Gln Ala Leu Tyr Gly Tyr Asn Pro  
 610 615 620  
 Pro Lys Gly Pro Lys Arg Pro Lys Met Val Ser Cys Asp Cys Cys Pro  
 625 630 635 640  
 Cys Phe Gly Ser Arg Lys Lys Tyr Lys Glu Lys Asn Asp Ala Asn Gly  
 645 650 655  
 Glu Ala Ala Ser Leu Lys Gly Met Asp Asp Asp Lys Glu Val Leu Met  
 660 665 670  
 Ser Gln Met Asn Phe Glu Lys Lys Phe Gly Gln Ser Ser Ile Phe Val  
 675 680 685  
 Thr Ser Thr Leu Met Glu Glu Gly Gly Val Pro Pro Ser Ser Ser Pro  
 690 695 700



Ala Ala Leu Leu Lys Glu Ala Ile His Val Ile Ser Cys Gly Tyr Glu  
 705 710 715 720  
 Asp Lys Thr Glu Trp Gly Leu Glu Leu Gly Trp Ile Tyr Gly Ser Ile  
 725 730 735  
 Thr Glu Asp Ile Leu Thr Gly Phe Lys Met His Cys Arg Gly Trp Arg  
 740 745 750  
 Ser Ile Tyr Cys Met Pro Lys Arg Ala Ala Phe Lys Gly Thr Ala Pro  
 755 760 765  
 Ile Asn Leu Ser Asp Arg Leu Asn Gln Val Leu Arg Trp Ala Leu Gly  
 770 775 780  
 Ser Ile Glu Ile Phe Phe Ser His His Cys Pro Leu Trp Tyr Gly Phe  
 785 790 795 800  
 Lys Glu Lys Lys Leu Lys Trp Leu Glu Arg Phe Ala Tyr Ala Asn Thr  
 805 810 815  
 Thr Val Tyr Pro Phe Thr Ser Ile Pro Leu Val Ala Tyr Cys Ile Leu  
 820 825 830  
 Pro Ala Val Cys Leu Leu Thr Asp Lys Phe Ile Met Pro Pro Ile Ser  
 835 840 845  
 Thr Phe Ala Gly Leu Tyr Phe Val Ala Leu Phe Ser Ser Ile Ile Ala  
 850 855 860  
 Thr Gly Ile Leu Glu Leu Lys Trp Ser Gly Val Ser Ile Glu Glu Trp  
 865 870 875 880  
 Trp Arg Asn Glu Gln Phe Trp Val Ile Gly Gly Val Ser Ala His Leu  
 885 890 895  
 Phe Ala Val Ile Gln Gly Leu Leu Lys Val Leu Ala Gly Ile Asp Thr  
 900 905 910  
 Asn Phe Thr Val Thr Ser Lys Ala Thr Asp Asp Glu Glu Phe Gly Glu  
 915 920 925  
 Leu Tyr Thr Phe Lys Trp Thr Thr Leu Leu Ile Pro Pro Thr Thr Ile  
 930 935 940  
 Leu Ile Ile Asn Ile Val Gly Val Val Ala Gly Ile Ser Asp Ala Ile  
 945 950 955 960  
 Asn Asn Gly Tyr Gln Ser Trp Gly Pro Leu Phe Gly Lys Leu Phe Phe  
 965 970 975  
 Ser Phe Trp Val Ile Val His Leu Tyr Pro Phe Leu Lys Gly Leu Met  
 980 985 990  
 Gly Arg Gln Asn Arg Thr Pro Thr Ile Val Val Ile Trp Ser Val Leu  
 995 1000 1005  
 Leu Ala Ser Ile Phe Ser Leu Leu Trp Val Arg Ile Asp Pro Phe Val  
 1010 1015 1020

Leu Lys Thr Lys Gly Pro Asp Thr Lys Leu Cys Gly Ile Asn Cys  
 1025 1030 1035

<210> 15  
 <211> 2125  
 <212> DNA  
 <213> Glycine max

<400> 15  
 gccaaagctc agaagatgcc agaggaaggt tggacaatgc aggatggaac tccttggcct 60  
 ggaaataatc ctagggatca tccgggaatg attcaggtgt ttttaggtca tagtgggggg 120  
 ctggatacag atggaaatga gctgcctaga cttgtttatg tttctcgtga gaagcgacca 180  
 ggcttccaac atcacaagaa ggctggagct atgaatgctt tgattcgagt ttctgctgtc 240  
 ttgaccaatg gtgcataatc tctgaatgtg gattgtgac actatttcaa taatagcaaa 300  
 gccctcaaag aagccatgtg tttcatgatg gatcctgttc ttggaaagaa gacatgctat 360  
 gttcaatttc ctgagagatt tgacggcatt gacttgcacg atcgatatgc caatcgcaat 420  
 attgtgttct ttgatataca catgaaaggt caggatggtg ttcagggccc agtctatgtg 480  
 ggaactggtt gttgtttcaa taggcaagct ttgtatggtt atgatcctgt tttgactgag 540  
 gaagatttgg aacctaaccat tattgtaaag agttgttgcg gttctagaaa gaagggaaaag 600  
 ggtggcaata agaagtacag tgacaagaag aaggcgatgg gaagaactga atccactgta 660  
 cccatattta atatggaaga catagaggag ggtgttgaag gttatgatga tgaaaggaca 720  
 ctacttatgt ctcaaaagag cttggagaag cgttttggtc agtctccagt ttttattgct 780  
 gccactttca tggagcaggg tggcattcca ccttcaacga accctgcaac tcttcttaag 840  
 gaagcaatcc atgttatcag ctgtggttac gaagacaaga cagaatgggg caaagagatt 900  
 ggatggatct atggctctgt gacagaagat atcttgactg ggttcaagat gcatgctcgt 960  
 ggttggattt ccactctattg catgccacct cgcccagcat ttaagggttc tgctcctatc 1020  
 aatctttctg atcgtctcaa tcaggtgctt cggtgggcct tgggttcaat tgagatcttt 1080  
 ctaagcaggg attgtccctt gtggtatggc tacaatggga agttgaagcc tctgatgagg 1140  
 cttgcttata ttaacacccat tgtctaccgg tttacctcaa tcccattgat tgcttactgt 1200  
 acgcttcctg ctttttgtct tctcacaat aaatttatta ttcttgagat aagcaacttt 1260  
 gccagtatgt ggttcattct tctctttgtc tccattttta ccacttcaat tcttgagctt 1320  
 aggtggagtg gggtcagtat agaagactgg tggagaaatg aacagttctg ggttatcggt 1380  
 gggacatctg cgcactctctt tgctgtgttc caggggcttc taaaagtgtc tgctgggatc 1440  
 gatacaaat ttactgttac atcgaaggca tgggacgagg atggggactt tgccgagctt 1500  
 tatgtgttta aatggacatc acttctcatc cctcctacaa cagtgtctat tgtgaatttg 1560  
 gttgggattg tggctggtgt atcctatgcc ataaacagtg gttaccagtc ttgggggtcca 1620  
 ctatttggca agctgttctt tgctatctgg gtcattgccc atctataccc attcttgaag 1680  
 ggtctcttgg gcaggcaaaa tcgtaccca accattgtta ttggttggtc cgttcttctt 1740  
 gcttcaatat tctccttgct gtgggtgagg attgatccct tcacctctga ctccaacaaa 1800  
 ttaaccaatg gtcaatgtgg catcaactgt tagttctctt gtatgattca ttttgtgttg 1860  
 ttattccctt ttgcttgagg atacacaagg ttgctgtcgt gtatatagca agaattttca 1920  
 gectatcaaa gttgtctgga ggattgaacc cctgaaatag atgggaatgt accctctctg 1980  
 tttctattat ttatctacat gttccttaca agaatagtca gtagtaatgt tgaggtgtat 2040  
 gttatatttt tccccacag aatataaatt tgttcatgcg aatatttaat gaaagccaac 2100  
 aaggtcctgt gttgttttgt tcttt 2125

<210> 16  
 <211> 610  
 <212> PRT  
 <213> Glycine max

<400> 16  
 Ala Lys Ala Gln Lys Met Pro Glu Glu Gly Trp Thr Met Gln Asp Gly  
 1 5 10 15  
 Thr Pro Trp Pro Gly Asn Asn Pro Arg Asp His Pro Gly Met Ile Gln  
 20 25 30

Val Phe Leu Gly His Ser Gly Gly Leu Asp Thr Asp Gly Asn Glu Leu  
                   35                                  40                                  45  
 Pro Arg Leu Val Tyr Val Ser Arg Glu Lys Arg Pro Gly Phe Gln His  
           50                                  55                                  60  
 His Lys Lys Ala Gly Ala Met Asn Ala Leu Ile Arg Val Ser Ala Val  
   65                                  70                                  75                                  80  
 Leu Thr Asn Gly Ala Tyr Leu Leu Asn Val Asp Cys Asp His Tyr Phe  
                                   85                                  90                                  95  
 Asn Asn Ser Lys Ala Leu Lys Glu Ala Met Cys Phe Met Met Asp Pro  
                   100                                  105                                  110  
 Val Leu Gly Lys Lys Thr Cys Tyr Val Gln Phe Pro Gln Arg Phe Asp  
                   115                                  120                                  125  
 Gly Ile Asp Leu His Asp Arg Tyr Ala Asn Arg Asn Ile Val Phe Phe  
   130                                  135                                  140  
 Asp Ile Asn Met Lys Gly Gln Asp Gly Val Gln Gly Pro Val Tyr Val  
  145                                  150                                  155                                  160  
 Gly Thr Gly Cys Cys Phe Asn Arg Gln Ala Leu Tyr Gly Tyr Asp Pro  
                                   165                                  170                                  175  
 Val Leu Thr Glu Glu Asp Leu Glu Pro Asn Ile Ile Val Lys Ser Cys  
                   180                                  185                                  190  
 Cys Gly Ser Arg Lys Lys Gly Lys Gly Gly Asn Lys Lys Tyr Ser Asp  
           195                                  200                                  205  
 Lys Lys Lys Ala Met Gly Arg Thr Glu Ser Thr Val Pro Ile Phe Asn  
   210                                  215                                  220  
 Met Glu Asp Ile Glu Glu Gly Val Glu Gly Tyr Asp Asp Glu Arg Thr  
  225                                  230                                  235                                  240  
 Leu Leu Met Ser Gln Lys Ser Leu Glu Lys Arg Phe Gly Gln Ser Pro  
                   245                                  250                                  255  
 Val Phe Ile Ala Ala Thr Phe Met Glu Gln Gly Gly Ile Pro Pro Ser  
                   260                                  265                                  270  
 Thr Asn Pro Ala Thr Leu Leu Lys Glu Ala Ile His Val Ile Ser Cys  
           275                                  280                                  285  
 Gly Tyr Glu Asp Lys Thr Glu Trp Gly Lys Glu Ile Gly Trp Ile Tyr  
   290                                  295                                  300  
 Gly Ser Val Thr Glu Asp Ile Leu Thr Gly Phe Lys Met His Ala Arg  
  305                                  310                                  315                                  320  
 Gly Trp Ile Ser Ile Tyr Cys Met Pro Pro Arg Pro Ala Phe Lys Gly  
                   325                                  330                                  335  
 Ser Ala Pro Ile Asn Leu Ser Asp Arg Leu Asn Gln Val Leu Arg Trp  
                   340                                  345                                  350

Ala Leu Gly Ser Ile Glu Ile Phe Leu Ser Arg His Cys Pro Leu Trp  
 355 360 365  
 Tyr Gly Tyr Asn Gly Lys Leu Lys Pro Leu Met Arg Leu Ala Tyr Ile  
 370 375 380  
 Asn Thr Ile Val Tyr Pro Phe Thr Ser Ile Pro Leu Ile Ala Tyr Cys  
 385 390 395 400  
 Thr Leu Pro Ala Phe Cys Leu Leu Thr Asn Lys Phe Ile Ile Pro Glu  
 405 410 415  
 Ile Ser Asn Phe Ala Ser Met Trp Phe Ile Leu Leu Phe Val Ser Ile  
 420 425 430  
 Phe Thr Thr Ser Ile Leu Glu Leu Arg Trp Ser Gly Val Ser Ile Glu  
 435 440 445  
 Asp Trp Trp Arg Asn Glu Gln Phe Trp Val Ile Gly Gly Thr Ser Ala  
 450 455 460  
 His Leu Phe Ala Val Phe Gln Gly Leu Leu Lys Val Leu Ala Gly Ile  
 465 470 475 480  
 Asp Thr Asn Phe Thr Val Thr Ser Lys Ala Ser Asp Glu Asp Gly Asp  
 485 490 495  
 Phe Ala Glu Leu Tyr Val Phe Lys Trp Thr Ser Leu Leu Ile Pro Pro  
 500 505 510  
 Thr Thr Val Leu Ile Val Asn Leu Val Gly Ile Val Ala Gly Val Ser  
 515 520 525  
 Tyr Ala Ile Asn Ser Gly Tyr Gln Ser Trp Gly Pro Leu Phe Gly Lys  
 530 535 540  
 Leu Phe Phe Ala Ile Trp Val Ile Ala His Leu Tyr Pro Phe Leu Lys  
 545 550 555 560  
 Gly Leu Leu Gly Arg Gln Asn Arg Thr Pro Thr Ile Val Ile Val Trp  
 565 570 575  
 Ser Val Leu Leu Ala Ser Ile Phe Ser Leu Leu Trp Val Arg Ile Asp  
 580 585 590  
 Pro Phe Thr Ser Asp Ser Asn Lys Leu Thr Asn Gly Gln Cys Gly Ile  
 595 600 605  
 Asn Cys  
 610  
 <210> 17  
 <211> 2890  
 <212> DNA  
 <213> Glycine max  
 <400> 17  
 gcacgagctc caccagtgta atgatgcata tggcttgtgg ttgacatcag tcattctgtga 60  
 aatatggttt gctgtatcat ggataatgga tcagtttcca aaatggtacc caatacagcg 120  
 agaaacatac cttgatcgct tgtcactcag gtatgaaaaa gaagggaagc catctgagtt 180

```

gtccagtgtgta gacgtctttg tcagtactgt tgatcccatg aaggaacctc cactgattac 240
agcaaacact gttctatcta tccttgctgt tgattatcca gttgataaag ttgcatgcta 300
tgtctcagat gatggtgctg ctatgcttac ttttgaagca ctgtctgaga catctgaatt 360
tgctaggaga tgggttccat tttgtaagaa atacaatatt gagccccggg caccagaatg 420
gtacttttggg cagaagatgg actatctgaa aaataaagta caccagcat ttgtcaggga 480
aaggagagca atgaagaggg attatgaaga atttaagggt aggattaaca gtttgggtggc 540
aacagcacaa aagggttcctg aggatggatg gaccatgcaa gatgggactc cttggcctgg 600
aaataatgtg agggatcatc ctggcatgat tcagggtcttc cttgggcagg atggtgttcg 660
tgatggtgaa ggaatgagc taccgcgtt ggtctacgtt tctagagaaa agaggccagg 720
gtttgatcac caaaaaagg ctggtgcaat gaatgctctg gtacgggctt cagcaattat 780
cactaatgca ccctatcttc tgaatgttga ttgtgatcac tacattaaca atagcaaggc 840
actagagaa gctatgtgtt ttatgatgga tcctcaacta gggaaaaagg tttgctatgt 900
gcaatttcct cagcgatttg atggaattga tagacatgat agatattcaa acagaaatgt 960
tgtatttttc gatattaaca tgaaggatt ggatgggata caagggtccaa tatatgtcgg 1020
aactggatgt gttttcagaa ggtacgcact ttatggatat gatgcacctg ccaagaagaa 1080
accaccgagc aaaacttgta actgttggcc aaagtgggtg tgccatgtt gtggctctag 1140
aaagaaaaag aatgccataa gtaagaagga gaaaaagagg aaggtgaagc acagtgaagc 1200
atcaaagcag atacatgcac ttgaaaatat tgaggcgggg aatgaaggaa ccaacaatga 1260
gaagacatcc aatctgactc aaacaaagtt ggagaagagg tttggacagt ctccagtatt 1320
tgtagcctcc acacttttgg atgatggtg agttccacat ggcgtgagtc ctgcatcact 1380
tttaaaagaa gccatccagg tcatcagttg tggttatgaa gacaaaacag aatggggaaa 1440
agaagttggg tggtatatg gttctgtgac agaggatatc ttgactggat ttaaaatgca 1500
ttgccatggt tggcggtctg tgtattgcat tcctaagcgg cctgcattta aggggtctgc 1560
gcctatcaac ctttcagatc gtctgcacca agttcttcgg tgggtctctg ggtctgttga 1620
gatttttttc agcagacatt gtccaatctg gtatggctat ggtgggtggat tgaattgtt 1680
ggaacgattt tcctacatta actcggctcg atatccctgg acttccctcc cattgcttgt 1740
ctactgtact ctaccagcca tatgccttct gactgaaaaa tttatcgtac ccgagattag 1800
caactatgcc agtcttgtgt tcatggcctt cttcatatcc attgcagcaa ctggcattcg 1860
tgagatgcaa tggggcggtg ttagcataga cgactgggtg aggaacgaac agttttgggt 1920
gatcggaggt gtttcttccc atctatttgc cctatttcag ggtttactga aggtcttggc 1980
tggtgtgaac acaaacttca ctgtgacctc aaaagcagca gatgatggag aattctcaga 2040
actctacata ttcaagtgga catcactctt gatccctcca atgactttac ttatcatgaa 2100
tattgtcggg gtggttgcg ggatctcaga tgccatcaac aatggttatg actcatggg 2160
acctctcttt ggtagattgt tctttgcatt gtgggtgatc ctccatctct accccttctt 2220
gaaggggttg cttggaaaac aagatagaat gccaacatt atattggtt ggtcaatct 2280
tctggcctcc atcttgactc tcatgtgggt cagaattaac ccgtttgtgt caagagacgg 2340
ccccgtgtta gaaatttgtg gattgaattg cgacgagtcg tgaataaaga aaagctgaag 2400
aagaagggtt agttattttt cagctacact gcagtcatgt tgaagaatgc agccagcaca 2460
tgcttcacaa agttgcacga attttcggat ggaagtttta ttttccgggt gtgtagagat 2520
taaagagagg aaggggaggg ggctgacaca ttgttacctt gtaatagggt ttttccattt 2580
attccttgat tatattttct gtgggtttta gtgttattct ctttccagtt tcatgtatta 2640
taagaaagag gcattgaatg ataaattatt cctcttcaa aatgggggat cctcagctc 2700
aagaaattac ttggtcatat ttttaggtat ggggtctgtt ctgtttaaaa ccatttgtaa 2760
taccgtcaaa actatggata ttcttgttcc tcagatgtgt ttttgtgttt tattatttaa 2820
cactcaggaa ccttttggtt tgattcaatt attcaatgtt tggatggcac taaaaaaaaa 2880
aaaaaaaaaa 2890

```

<210> 18  
 <211> 793  
 <212> PRT  
 <213> Glycine max

<400> 18  
 His Glu Leu His Pro Val Asn Asp Ala Tyr Gly Leu Trp Leu Thr Ser  
 1 5 10 15  
 Val Ile Cys Glu Ile Trp Phe Ala Val Ser Trp Ile Met Asp Gln Phe  
 20 25 30

Pro Lys Trp Tyr Pro Ile Gln Arg Glu Thr Tyr Leu Asp Arg Leu Ser  
                   35                                  40                                  45  
 Leu Arg Tyr Glu Lys Glu Gly Lys Pro Ser Glu Leu Ser Ser Val Asp  
                   50                                  55                                  60  
 Val Phe Val Ser Thr Val Asp Pro Met Lys Glu Pro Pro Leu Ile Thr  
                   65                                  70                                  75                                  80  
 Ala Asn Thr Val Leu Ser Ile Leu Ala Val Asp Tyr Pro Val Asp Lys  
                                   85                                  90                                  95  
 Val Ala Cys Tyr Val Ser Asp Asp Gly Ala Ala Met Leu Thr Phe Glu  
                                   100                                  105                                  110  
 Ala Leu Ser Glu Thr Ser Glu Phe Ala Arg Arg Trp Val Pro Phe Cys  
                                   115                                  120                                  125  
 Lys Lys Tyr Asn Ile Glu Pro Arg Ala Pro Glu Trp Tyr Phe Gly Gln  
                   130                                  135                                  140  
 Lys Met Asp Tyr Leu Lys Asn Lys Val His Pro Ala Phe Val Arg Glu  
                   145                                  150                                  155                                  160  
 Arg Arg Ala Met Lys Arg Asp Tyr Glu Glu Phe Lys Val Arg Ile Asn  
                                   165                                  170                                  175  
 Ser Leu Val Ala Thr Ala Gln Lys Val Pro Glu Asp Gly Trp Thr Met  
                                   180                                  185                                  190  
 Gln Asp Gly Thr Pro Trp Pro Gly Asn Asn Val Arg Asp His Pro Gly  
                   195                                  200                                  205  
 Met Ile Gln Val Phe Leu Gly Gln Asp Gly Val Arg Asp Val Glu Gly  
                   210                                  215                                  220  
 Asn Glu Leu Pro Arg Leu Val Tyr Val Ser Arg Glu Lys Arg Pro Gly  
                   225                                  230                                  235                                  240  
 Phe Asp His His Lys Lys Ala Gly Ala Met Asn Ala Leu Val Arg Ala  
                                   245                                  250                                  255  
 Ser Ala Ile Ile Thr Asn Ala Pro Tyr Leu Leu Asn Val Asp Cys Asp  
                                   260                                  265                                  270  
 His Tyr Ile Asn Asn Ser Lys Ala Leu Arg Glu Ala Met Cys Phe Met  
                   275                                  280                                  285  
 Met Asp Pro Gln Leu Gly Lys Lys Val Cys Tyr Val Gln Phe Pro Gln  
                   290                                  295                                  300  
 Arg Phe Asp Gly Ile Asp Arg His Asp Arg Tyr Ser Asn Arg Asn Val  
                   305                                  310                                  315                                  320  
 Val Phe Phe Asp Ile Asn Met Lys Gly Leu Asp Gly Ile Gln Gly Pro  
                                   325                                  330                                  335  
 Ile Tyr Val Gly Thr Gly Cys Val Phe Arg Arg Tyr Ala Leu Tyr Gly  
                   340                                  345                                  350

Tyr Asp Ala Pro Ala Lys Lys Lys Pro Pro Ser Lys Thr Cys Asn Cys  
 355 360 365  
 Trp Pro Lys Trp Cys Cys Leu Cys Cys Gly Ser Arg Lys Lys Lys Asn  
 370 375 380  
 Ala Asn Ser Lys Lys Glu Lys Lys Arg Lys Val Lys His Ser Glu Ala  
 385 390 395 400  
 Ser Lys Gln Ile His Ala Leu Glu Asn Ile Glu Ala Gly Asn Glu Gly  
 405 410 415  
 Thr Asn Asn Glu Lys Thr Ser Asn Leu Thr Gln Thr Lys Leu Glu Lys  
 420 425 430  
 Arg Phe Gly Gln Ser Pro Val Phe Val Ala Ser Thr Leu Leu Asp Asp  
 435 440 445  
 Gly Gly Val Pro His Gly Val Ser Pro Ala Ser Leu Leu Lys Glu Ala  
 450 455 460  
 Ile Gln Val Ile Ser Cys Gly Tyr Glu Asp Lys Thr Glu Trp Gly Lys  
 465 470 475 480  
 Glu Val Gly Trp Ile Tyr Gly Ser Val Thr Glu Asp Ile Leu Thr Gly  
 485 490 495  
 Phe Lys Met His Cys His Gly Trp Arg Ser Val Tyr Cys Ile Pro Lys  
 500 505 510  
 Arg Pro Ala Phe Lys Gly Ser Ala Pro Ile Asn Leu Ser Asp Arg Leu  
 515 520 525  
 His Gln Val Leu Arg Trp Ala Leu Gly Ser Val Glu Ile Phe Phe Ser  
 530 535 540  
 Arg His Cys Pro Ile Trp Tyr Gly Tyr Gly Gly Gly Leu Lys Leu Leu  
 545 550 555 560  
 Glu Arg Phe Ser Tyr Ile Asn Ser Val Val Tyr Pro Trp Thr Ser Leu  
 565 570 575  
 Pro Leu Leu Val Tyr Cys Thr Leu Pro Ala Ile Cys Leu Leu Thr Gly  
 580 585 590  
 Lys Phe Ile Val Pro Glu Ile Ser Asn Tyr Ala Ser Leu Val Phe Met  
 595 600 605  
 Ala Leu Phe Ile Ser Ile Ala Ala Thr Gly Ile Leu Glu Met Gln Trp  
 610 615 620  
 Gly Gly Val Ser Ile Asp Asp Trp Trp Arg Asn Glu Gln Phe Trp Val  
 625 630 635 640  
 Ile Gly Gly Val Ser Ser His Leu Phe Ala Leu Phe Gln Gly Leu Leu  
 645 650 655  
 Lys Val Leu Ala Gly Val Asn Thr Asn Phe Thr Val Thr Ser Lys Ala  
 660 665 670

Ala Asp Asp Gly Glu Phe Ser Glu Leu Tyr Ile Phe Lys Trp Thr Ser  
675 680 685

Leu Leu Ile Pro Pro Met Thr Leu Leu Ile Met Asn Ile Val Gly Val  
690 695 700

Val Val Gly Ile Ser Asp Ala Ile Asn Asn Gly Tyr Asp Ser Trp Gly  
705 710 715 720

Pro Leu Phe Gly Arg Leu Phe Phe Ala Leu Trp Val Ile Leu His Leu  
725 730 735

Tyr Pro Phe Leu Lys Gly Leu Leu Gly Lys Gln Asp Arg Met Pro Thr  
740 745 750

Ile Ile Leu Val Trp Ser Ile Leu Leu Ala Ser Ile Leu Thr Leu Met  
755 760 765

Trp Val Arg Ile Asn Pro Phe Val Ser Arg Asp Gly Pro Val Leu Glu  
770 775 780

Ile Cys Gly Leu Asn Cys Asp Glu Ser  
785 790

<210> 19  
<211> 1742  
<212> DNA  
<213> Triticum aestivum

<220>  
<221> unsure  
<222> (9)

<220>  
<221> unsure  
<222> (271)

<400> 19  
>wheatcong aagctatgtg cttcctaatt gatccaaacc taggtccgca agtctgttat 60  
gtgcagttcc cacaaagggt tgatgggtatt gataggaatg atcgatatgc aaacaggaac 120  
actgtctttt ttgatattaa cttgaggggc cttgacggca ttcaaggacc agtttatgtg 180  
ggaactgggt gtgttttcaa cagaacggct atctatgggt atgagccccc aattaaggcg 240  
aagaagccag gtttcttggc atcattatgt nggggcaaga agaaggcaag caagtcaaag 300  
aaaaggagct cagataagaa aaagtgaac aagcatgtgg acagttctgt tccagtattc 360  
aatctcgaag acatagagga ggggtgttgaa ggtgctgggt ttgatgatga gaaatcagtt 420  
ctcatgtctc aaatgagctt agagaagaga ttggccagt cagcagcatt tgttgccctc 480  
actctgatgg aatatgggtg tgttcctcag tcgtccactc cagaatctct tttgaaagaa 540  
gctatccatg tcataagttg tggctatgag gacaagtctg aatggggaac tgagattggg 600  
tggatctatg gatctgtcac agaagatatt ctaactggat tcaagatgca cgcaagaggc 660  
tggcggttaa tctattgcat gccaagcgc ccagctttca agggatctgc ccccatcaat 720  
ctttcagatc gtctgaatca agtgctgcgg tgggctcttg gttctgttga aattcttttc 780  
agccggcatt gccccttatg gtatggctac ggagggcgcc tcaagttcct ggagagattc 840  
gcttacatca acaccacat ttaccacta acctctctcc cgcttctagt ctattgtata 900  
ttgcctgcta tctgtctgct cactggaaag ttcattcatg cagagattag caacttggcc 960  
agtatctggg tcattgcgct cttcctttca attttcgcca ctgggtatcct tgagatgagg 1020  
tggagtgggt ttggcattga cgagtgggtg aggaatgaac agttctgggt cattggagg 1080  
atctctgccc atctgtttgc cgtctttcag ggtcttctga aggtgcttgc aggtatcgac 1140  
accaacttca ctgtcacctc aaaggctaag gatgaagaag gcgactttgc tgagctctac 1200  
atgttcaagt ggacgacgct tcttatccct ccgacgacca ttttgatcat taacatgggc 1260  
ggtgtcgttg ctggtacctc ctacgccatc aacagtgggt accaatcatg ggggccgctc 1320



```

tttgggaagc tcttctttgc cttctgggtg attgttcaact tataccatt cctcaagggt 1380
cttatgggca ggcaaaaccg cacaccgacg attgtcatcg tctgggctgt cctcctcgct 1440
tctatcttct ccttgctgtg gggtcgtgtt gatccattca ctaccgctct cgctggccca 1500
aatatccaaa cctgtggcat caactgctag gaaagtggga gttttagag acagaaaata 1560
taacagtgat cgagcgacca cctgtggagc cagagaatat ttatgttggg gttgtgaatt 1620
actacgtttg agaaagtgtt caaaattgag aaaacacatt tgtaaataga tgtaatagac 1680
tatctaccgt tttcatgagg ttaagctctt cttttttgga aaaaaaaaaa aaaaaaaaaa 1740
aa

```

```

<210> 20
<211> 506
<212> PRT
<213> Triticum aestivum

```

```

<220>
<221> UNSURE
<222> (88)

```

```

<400> 20
Glu Ala Met Cys Phe Leu Met Asp Pro Asn Leu Gly Pro Gln Val Cys
 1          5          10          15
Tyr Val Gln Phe Pro Gln Arg Phe Asp Gly Ile Asp Arg Asn Asp Arg
          20          25          30
Tyr Ala Asn Arg Asn Thr Val Phe Phe Asp Ile Asn Leu Arg Gly Leu
          35          40          45
Asp Gly Ile Gln Gly Pro Val Tyr Val Gly Thr Gly Cys Val Phe Asn
          50          55          60
Arg Thr Ala Ile Tyr Gly Tyr Glu Pro Pro Ile Lys Ala Lys Lys Pro
          65          70          75          80
Gly Phe Leu Ala Ser Leu Cys Xaa Gly Lys Lys Lys Ala Ser Lys Ser
          85          90          95
Lys Lys Arg Ser Ser Asp Lys Lys Lys Ser Asn Lys His Val Asp Ser
          100          105          110
Ser Val Pro Val Phe Asn Leu Glu Asp Ile Glu Glu Gly Val Glu Gly
          115          120          125
Ala Gly Phe Asp Asp Glu Lys Ser Val Leu Met Ser Gln Met Ser Leu
          130          135          140
Glu Lys Arg Phe Gly Gln Ser Ala Ala Phe Val Ala Ser Thr Leu Met
          145          150          155          160
Glu Tyr Gly Gly Val Pro Gln Ser Ser Thr Pro Glu Ser Leu Leu Lys
          165          170          175
Glu Ala Ile His Val Ile Ser Cys Gly Tyr Glu Asp Lys Ser Glu Trp
          180          185          190
Gly Thr Glu Ile Gly Trp Ile Tyr Gly Ser Val Thr Glu Asp Ile Leu
          195          200          205
Thr Gly Phe Lys Met His Ala Arg Gly Trp Arg Ser Ile Tyr Cys Met
          210          215          220

```

Pro Lys Arg Pro Ala Phe Lys Gly Ser Ala Pro Ile Asn Leu Ser Asp  
 225 230 235 240  
 Arg Leu Asn Gln Val Leu Arg Trp Ala Leu Gly Ser Val Glu Ile Leu  
 245 250 255  
 Phe Ser Arg His Cys Pro Leu Trp Tyr Gly Tyr Gly Gly Arg Leu Lys  
 260 265 270  
 Phe Leu Glu Arg Phe Ala Tyr Ile Asn Thr Thr Ile Tyr Pro Leu Thr  
 275 280 285  
 Ser Leu Pro Leu Leu Val Tyr Cys Ile Leu Pro Ala Ile Cys Leu Leu  
 290 295 300  
 Thr Gly Lys Phe Ile Met Pro Glu Ile Ser Asn Leu Ala Ser Ile Trp  
 305 310 315 320  
 Phe Ile Ala Leu Phe Leu Ser Ile Phe Ala Thr Gly Ile Leu Glu Met  
 325 330 335  
 Arg Trp Ser Gly Val Gly Ile Asp Glu Trp Trp Arg Asn Glu Gln Phe  
 340 345 350  
 Trp Val Ile Gly Gly Ile Ser Ala His Leu Phe Ala Val Phe Gln Gly  
 355 360 365  
 Leu Leu Lys Val Leu Ala Gly Ile Asp Thr Asn Phe Thr Val Thr Ser  
 370 375 380  
 Lys Ala Asn Asp Glu Glu Gly Asp Phe Ala Glu Leu Tyr Met Phe Lys  
 385 390 395 400  
 Trp Thr Thr Leu Leu Ile Pro Pro Thr Thr Ile Leu Ile Ile Asn Met  
 405 410 415  
 Val Gly Val Val Ala Gly Thr Ser Tyr Ala Ile Asn Ser Gly Tyr Gln  
 420 425 430  
 Ser Trp Gly Pro Leu Phe Gly Lys Leu Phe Phe Ala Phe Trp Val Ile  
 435 440 445  
 Val His Leu Tyr Pro Phe Leu Lys Gly Leu Met Gly Arg Gln Asn Arg  
 450 455 460  
 Thr Pro Thr Ile Val Ile Val Trp Ala Val Leu Leu Ala Ser Ile Phe  
 465 470 475 480  
 Ser Leu Leu Trp Val Arg Val Asp Pro Phe Thr Thr Arg Leu Ala Gly  
 485 490 495  
 Pro Asn Ile Gln Thr Cys Gly Ile Asn Cys  
 500 505

<210> 21  
 <211> 1029  
 <212> DNA  
 <213> Triticum aestivum

```

<400> 21
gcacgagccg ctctcacca acggccagat ggttgatgac atcccgccgg agcagcacgc 60
gctcgtgccg tcctacatga gcggcggcgg cggcgggggc aagaggatcc acccgctccc 120
tttcgcagat cccaaccttc cagtgc aacc gagatccatg gacccgtcca aggatctggc 180
cgcttacgga tatggcagcg tggcctggaa ggagagaatg gagggctgga agcagaagca 240
ggagcgccg cagcatgtca ggagcgaggg tggcggtgat tgggatggcg acgatgcaga 300
tctgccacta atggatgaag ctaggcagcc attgtccaga aaagtcccta tatcatcaag 360
ccgaattaat ccctacagga tgattatcgt tatccggttg gtggttttgg gtttcttctt 420
ccactaccga gtgatgcac cggcgaaaga tgcatttgca ttgtggctca tatctgtaat 480
ctgtgaaatc tggtttgca tgcctgtat tcttgatcag ttcccaaagt ggtttccaat 540
cgagagagag acttacctgg accgtttgtc actaagggtt gacaaggaag gtcaaccctc 600
tcagcttgct ccaatcgact tctttgtcag tacgggttgat ccacaaagg aacctccctt 660
ggtcacagcg aacactgtcc ttcccatcct ttctgtggat tatccggttg agaaggtctc 720
ctgctatggt tctgatgatg gtgctgcaat gcttacgttt gaagcattgt ctgaaacatc 780
tgaatttgca aagaaatggg ttcttttcag caaaaagttt aatatcgagc ctctgtctcc 840
tgagtggtag ttccaacaga agatagacta cctgaaagac aaggttgctg cttcatttgt 900
tagggagagg agggcgatga agagagaata cgaggaattc aaggtaagga tcaatgcctt 960
ggttgcaaaa gcccaaaagg ttcttgagga aggatggaca atgcaagatg gaagcccctg 1020
gcctggaaa 1029

```

```

<210> 22
<211> 340
<212> PRT
<213> Triticum aestivum

```

```

<400> 22
Pro Leu Leu Thr Asn Gly Gln Met Val Asp Asp Ile Pro Pro Glu Gln
1 5 10 15
His Ala Leu Val Pro Ser Tyr Met Ser Gly Gly Gly Gly Gly Gly Lys
20 25 30
Arg Ile His Pro Leu Pro Phe Ala Asp Pro Asn Leu Pro Val Gln Pro
35 40 45
Arg Ser Met Asp Pro Ser Lys Asp Leu Ala Ala Tyr Gly Tyr Gly Ser
50 55 60
Val Ala Trp Lys Glu Arg Met Glu Gly Trp Lys Gln Lys Gln Glu Arg
65 70 75 80
Leu Gln His Val Arg Ser Glu Gly Gly Gly Asp Trp Asp Gly Asp Asp
85 90 95
Ala Asp Leu Pro Leu Met Asp Glu Ala Arg Gln Pro Leu Ser Arg Lys
100 105 110
Val Pro Ile Ser Ser Ser Arg Ile Asn Pro Tyr Arg Met Ile Ile Val
115 120 125
Ile Arg Leu Val Val Leu Gly Phe Phe Phe His Tyr Arg Val Met His
130 135 140
Pro Ala Lys Asp Ala Phe Ala Leu Trp Leu Ile Ser Val Ile Cys Glu
145 150 155 160
Ile Trp Phe Ala Met Ser Cys Ile Leu Asp Gln Phe Pro Lys Trp Phe
165 170 175

```

Pro Ile Glu Arg Glu Thr Tyr Leu Asp Arg Leu Ser Leu Arg Phe Asp  
                     180                    185                    190  
 Lys Glu Gly Gln Pro Ser Gln Leu Ala Pro Ile Asp Phe Phe Val Ser  
                     195                    200                    205  
 Thr Val Asp Pro Thr Lys Glu Pro Pro Leu Val Thr Ala Asn Thr Val  
                     210                    215                    220  
 Leu Ser Ile Leu Ser Val Asp Tyr Pro Val Glu Lys Val Ser Cys Tyr  
                     225                    230                    235                    240  
 Val Ser Asp Asp Gly Ala Ala Met Leu Thr Phe Glu Ala Leu Ser Glu  
                     245                    250                    255  
 Thr Ser Glu Phe Ala Lys Lys Trp Val Pro Phe Ser Lys Lys Phe Asn  
                     260                    265                    270  
 Ile Glu Pro Arg Ala Pro Glu Trp Tyr Phe Gln Gln Lys Ile Asp Tyr  
                     275                    280                    285  
 Leu Lys Asp Lys Val Ala Ala Ser Phe Val Arg Glu Arg Arg Ala Met  
                     290                    295                    300  
 Lys Arg Glu Tyr Glu Glu Phe Lys Val Arg Ile Asn Ala Leu Val Ala  
                     305                    310                    315                    320  
 Lys Ala Gln Lys Val Pro Glu Glu Gly Trp Thr Met Gln Asp Gly Ser  
                     325                    330                    335  
 Pro Trp Pro Gly  
                     340

<210> 23  
 <211> 1081  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 23  
 Met Glu Ala Ser Ala Gly Leu Val Ala Gly Ser Tyr Arg Arg Asn Glu  
   1                    5                    10                    15  
 Leu Val Arg Ile Arg His Glu Ser Asp Gly Gly Thr Lys Pro Leu Lys  
                     20                    25                    30  
 Asn Met Asn Gly Gln Ile Cys Gln Ile Cys Gly Asp Asp Val Gly Leu  
                     35                    40                    45  
 Ala Glu Thr Gly Asp Val Phe Val Ala Cys Asn Glu Cys Ala Phe Pro  
                     50                    55                    60  
 Val Cys Arg Pro Cys Tyr Glu Tyr Glu Arg Lys Asp Gly Thr Gln Cys  
                     65                    70                    75                    80  
 Cys Pro Gln Cys Lys Thr Arg Phe Arg Arg His Arg Gly Ser Pro Arg  
                     85                    90                    95  
 Val Glu Gly Asp Glu Asp Glu Asp Asp Val Asp Asp Ile Glu Asn Glu  
                     100                    105                    110

Phe Asn Tyr Ala Gln Gly Ala Asn Lys Ala Arg His Gln Arg His Gly  
 115 120 125  
 Glu Glu Phe Ser Ser Ser Ser Arg His Glu Ser Gln Pro Ile Pro Leu  
 130 135 140  
 Leu Thr His Gly His Thr Val Ser Gly Glu Ile Arg Thr Pro Asp Thr  
 145 150 155 160  
 Gln Ser Val Arg Thr Thr Ser Gly Pro Leu Gly Pro Ser Asp Arg Asn  
 165 170 175  
 Ala Ile Ser Ser Pro Tyr Ile Asp Pro Arg Gln Pro Val Pro Val Arg  
 180 185 190  
 Ile Val Asp Pro Ser Lys Asp Leu Asn Ser Tyr Gly Leu Gly Asn Val  
 195 200 205  
 Asp Trp Lys Glu Arg Val Glu Gly Trp Lys Leu Lys Gln Glu Lys Asn  
 210 215 220  
 Met Leu Gln Met Thr Gly Lys Tyr His Glu Gly Lys Gly Gly Glu Ile  
 225 230 235 240  
 Glu Gly Thr Gly Ser Asn Gly Glu Glu Leu Gln Met Ala Asp Asp Thr  
 245 250 255  
 Arg Leu Pro Met Ser Arg Val Val Pro Ile Pro Ser Ser Arg Leu Thr  
 260 265 270  
 Pro Tyr Arg Val Val Ile Ile Leu Arg Leu Ile Ile Leu Cys Phe Phe  
 275 280 285  
 Leu Gln Tyr Arg Thr Thr His Pro Val Lys Asn Ala Tyr Pro Leu Trp  
 290 295 300  
 Leu Thr Ser Val Ile Cys Glu Ile Trp Phe Ala Phe Ser Trp Leu Leu  
 305 310 315 320  
 Asp Gln Phe Pro Lys Trp Tyr Pro Ile Asn Arg Glu Thr Tyr Leu Asp  
 325 330 335  
 Arg Leu Ala Ile Arg Tyr Asp Arg Asp Gly Glu Pro Ser Gln Leu Val  
 340 345 350  
 Pro Val Asp Val Phe Val Ser Thr Val Asp Pro Leu Lys Glu Pro Pro  
 355 360 365  
 Leu Val Thr Ala Asn Thr Val Leu Ser Ile Leu Ser Val Asp Tyr Pro  
 370 375 380  
 Val Asp Lys Val Ala Cys Tyr Val Ser Asp Asp Gly Ser Ala Met Leu  
 385 390 395 400  
 Thr Phe Glu Ser Leu Ser Glu Thr Ala Glu Phe Ala Lys Lys Trp Val  
 405 410 415  
 Pro Phe Cys Lys Lys Phe Asn Ile Glu Pro Arg Ala Pro Glu Phe Tyr  
 420 425 430

Phe Ala Gln Lys Ile Asp Tyr Leu Lys Asp Lys Ile Gln Pro Ser Phe  
 435 440 445  
 Val Lys Glu Arg Arg Ala Met Lys Arg Glu Tyr Glu Glu Phe Lys Val  
 450 455 460  
 Arg Ile Asn Ala Leu Val Ala Lys Ala Gln Lys Ile Pro Glu Glu Gly  
 465 470 475 480  
 Trp Thr Met Gln Asp Gly Thr Pro Trp Pro Gly Asn Asn Thr Arg Asp  
 485 490 495  
 His Pro Gly Met Ile Gln Val Phe Leu Gly His Ser Gly Gly Leu Asp  
 500 505 510  
 Thr Asp Gly Asn Glu Leu Pro Arg Leu Ile Tyr Val Ser Arg Glu Lys  
 515 520 525  
 Arg Pro Gly Phe Gln His His Lys Lys Ala Gly Ala Met Asn Ala Leu  
 530 535 540  
 Ile Arg Val Ser Ala Val Leu Thr Asn Gly Ala Tyr Leu Leu Asn Val  
 545 550 555 560  
 Asp Cys Asp His Tyr Phe Asn Asn Ser Lys Ala Ile Lys Glu Ala Met  
 565 570 575  
 Cys Phe Met Met Asp Pro Ala Ile Gly Lys Lys Cys Cys Tyr Val Gln  
 580 585 590  
 Phe Pro Gln Arg Phe Asp Gly Ile Asp Leu His Asp Arg Tyr Ala Asn  
 595 600 605  
 Arg Asn Ile Val Phe Phe Asp Ile Asn Met Lys Gly Leu Asp Gly Ile  
 610 615 620  
 Gln Gly Pro Val Tyr Val Gly Thr Gly Cys Cys Phe Asn Arg Gln Ala  
 625 630 635 640  
 Leu Tyr Gly Tyr Asp Pro Val Leu Thr Glu Glu Asp Leu Glu Pro Asn  
 645 650 655  
 Ile Ile Val Lys Ser Cys Cys Gly Ser Arg Lys Lys Gly Lys Ser Ser  
 660 665 670  
 Lys Lys Tyr Asn Tyr Glu Lys Arg Arg Gly Ile Asn Arg Ser Asp Ser  
 675 680 685  
 Asn Ala Pro Leu Phe Asn Met Glu Asp Ile Asp Glu Gly Phe Glu Gly  
 690 695 700  
 Tyr Asp Asp Glu Arg Ser Ile Leu Met Ser Gln Arg Ser Val Glu Lys  
 705 710 715 720  
 Arg Phe Gly Gln Ser Pro Val Phe Ile Ala Ala Thr Phe Met Glu Gln  
 725 730 735  
 Gly Gly Ile Pro Pro Thr Thr Asn Pro Ala Thr Leu Leu Lys Glu Ala  
 740 745 750

Ile His Val Ile Ser Cys Gly Tyr Glu Asp Lys Thr Glu Trp Gly Lys  
 755 760 765  
 Glu Ile Gly Trp Ile Tyr Gly Ser Val Thr Glu Asp Ile Leu Thr Gly  
 770 775 780  
 Phe Lys Met His Ala Arg Gly Trp Ile Ser Ile Tyr Cys Asn Pro Pro  
 785 790 795 800  
 Arg Pro Ala Phe Lys Gly Ser Ala Pro Ile Asn Leu Ser Asp Arg Leu  
 805 810 815  
 Asn Gln Val Leu Arg Trp Ala Leu Gly Ser Ile Glu Ile Leu Leu Ser  
 820 825 830  
 Arg His Cys Pro Ile Trp Tyr Gly Tyr His Gly Arg Leu Arg Leu Leu  
 835 840 845  
 Glu Arg Ile Ala Tyr Ile Asn Thr Ile Val Tyr Pro Ile Thr Ser Ile  
 850 855 860  
 Pro Leu Ile Ala Tyr Cys Ile Leu Pro Ala Phe Cys Leu Ile Thr Asp  
 865 870 875 880  
 Arg Phe Ile Ile Pro Glu Ile Ser Asn Tyr Ala Ser Ile Trp Phe Ile  
 885 890 895  
 Leu Leu Phe Ile Ser Ile Ala Val Thr Gly Ile Leu Glu Leu Arg Trp  
 900 905 910  
 Ser Gly Val Ser Ile Glu Asp Trp Trp Arg Asn Glu Gln Phe Trp Val  
 915 920 925  
 Ile Gly Gly Thr Ser Ala His Leu Phe Ala Val Phe Gln Gly Leu Leu  
 930 935 940  
 Lys Val Leu Ala Gly Ile Asp Thr Asn Phe Thr Val Thr Ser Lys Ala  
 945 950 955 960  
 Thr Asp Glu Asp Gly Asp Phe Ala Glu Leu Tyr Ile Phe Lys Trp Thr  
 965 970 975  
 Ala Leu Leu Ile Pro Pro Thr Thr Val Leu Leu Val Asn Leu Ile Gly  
 980 985 990  
 Ile Val Ala Gly Val Ser Tyr Ala Val Asn Ser Gly Tyr Gln Ser Trp  
 995 1000 1005  
 Gly Pro Leu Phe Gly Lys Leu Phe Phe Ala Leu Trp Val Ile Ala His  
 1010 1015 1020  
 Leu Tyr Pro Phe Leu Lys Gly Leu Leu Gly Arg Gln Asn Arg Thr Pro  
 1025 1030 1035 1040  
 Thr Ile Val Ile Val Trp Ser Val Leu Leu Ala Ser Ile Phe Ser Leu  
 1045 1050 1055  
 Leu Trp Val Arg Ile Asn Pro Phe Val Asp Ala Asn Pro Asn Ala Asn  
 1060 1065 1070

Asn Phe Asn Gly Lys Gly Gly Val Phe  
 1075 1080

<210> 24  
 <211> 1084  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 24

Met Asn Thr Gly Gly Arg Leu Ile Ala Gly Ser His Asn Arg Asn Glu  
 1 5 10 15

Phe Val Leu Ile Asn Ala Asp Glu Ser Ala Arg Ile Arg Ser Val Gln  
 20 25 30

Glu Leu Ser Gly Gln Thr Cys Gln Ile Cys Gly Asp Glu Ile Glu Leu  
 35 40 45

Thr Val Ser Ser Glu Leu Phe Val Ala Cys Asn Glu Cys Ala Phe Pro  
 50 55 60

Val Cys Arg Pro Cys Tyr Glu Tyr Glu Arg Arg Glu Gly Asn Gln Ala  
 65 70 75 80

Cys Pro Gln Cys Lys Thr Arg Tyr Lys Arg Ile Lys Gly Ser Pro Arg  
 85 90 95

Val Asp Gly Asp Asp Glu Glu Glu Glu Asp Ile Asp Asp Leu Glu Tyr  
 100 105 110

Glu Phe Asp His Gly Met Asp Pro Glu His Ala Ala Glu Ala Ala Leu  
 115 120 125

Ser Ser Arg Leu Asn Thr Gly Arg Gly Gly Leu Asp Ser Ala Pro Pro  
 130 135 140

Gly Ser Gln Ile Pro Leu Leu Thr Tyr Cys Asp Glu Asp Ala Asp Met  
 145 150 155 160

Tyr Ser Asp Arg His Ala Leu Ile Val Pro Pro Ser Thr Gly Tyr Gly  
 165 170 175

Asn Arg Val Tyr Pro Ala Pro Phe Thr Asp Ser Ser Ala Pro Pro Gln  
 180 185 190

Ala Arg Ser Met Val Pro Gln Lys Asp Ile Ala Glu Tyr Gly Tyr Gly  
 195 200 205

Ser Val Ala Trp Lys Asp Arg Met Glu Val Trp Lys Arg Arg Gln Gly  
 210 215 220

Glu Lys Leu Gln Val Ile Lys His Glu Gly Gly Asn Asn Gly Arg Gly  
 225 230 235 240

Ser Asn Asp Asp Asp Glu Leu Asp Asp Pro Asp Met Pro Met Met Asp  
 245 250 255

Glu Gly Arg Gln Pro Leu Ser Arg Lys Leu Pro Ile Arg Ser Ser Arg  
 260 265 270



Ile Asn Pro Tyr Arg Met Leu Ile Leu Cys Arg Leu Ala Ile Leu Gly  
 275 280 285  
 Leu Phe Phe His Tyr Arg Ile Leu His Pro Val Asn Asp Ala Tyr Gly  
 290 295 300  
 Leu Trp Leu Thr Ser Val Ile Cys Glu Ile Trp Phe Ala Val Ser Trp  
 305 310 315 320  
 Ile Leu Asp Gln Phe Pro Lys Trp Tyr Pro Ile Glu Arg Glu Thr Tyr  
 325 330 335  
 Leu Asp Arg Leu Ser Leu Arg Tyr Glu Lys Glu Gly Lys Pro Ser Gly  
 340 345 350  
 Leu Ala Pro Val Asp Val Phe Val Ser Thr Val Asp Pro Leu Lys Glu  
 355 360 365  
 Pro Pro Leu Ile Thr Ala Asn Thr Val Leu Ser Ile Leu Ala Val Asp  
 370 375 380  
 Tyr Pro Val Asp Lys Val Ala Cys Tyr Val Ser Asp Asp Gly Ala Ala  
 385 390 395 400  
 Met Leu Thr Phe Glu Ala Leu Ser Asp Thr Ala Glu Phe Ala Arg Lys  
 405 410 415  
 Trp Val Pro Phe Cys Lys Lys Phe Asn Ile Glu Pro Arg Ala Pro Glu  
 420 425 430  
 Trp Tyr Phe Ser Gln Lys Met Asp Tyr Leu Lys Asn Lys Val His Pro  
 435 440 445  
 Ala Phe Val Arg Glu Arg Arg Ala Met Lys Arg Asp Tyr Glu Glu Phe  
 450 455 460  
 Lys Val Lys Ile Asn Ala Leu Val Ala Thr Ala Gln Lys Val Pro Glu  
 465 470 475 480  
 Glu Gly Trp Thr Met Gln Asp Gly Thr Pro Trp Pro Gly Asn Asn Val  
 485 490 495  
 Arg Asp His Pro Gly Met Ile Gln Val Phe Leu Gly His Ser Gly Val  
 500 505 510  
 Arg Asp Thr Asp Gly Asn Glu Leu Pro Arg Leu Val Tyr Val Ser Arg  
 515 520 525  
 Glu Lys Arg Pro Gly Phe Asp His His Lys Lys Ala Gly Ala Met Asn  
 530 535 540  
 Ser Leu Ile Arg Val Ser Ala Val Leu Ser Asn Ala Pro Tyr Leu Leu  
 545 550 555 560  
 Asn Val Asp Cys Asp His Tyr Ile Asn Asn Ser Lys Ala Ile Arg Glu  
 565 570 575  
 Ser Met Cys Phe Met Met Asp Pro Gln Ser Gly Lys Lys Val Cys Tyr  
 580 585 590

Val Gln Phe Pro Gln Arg Phe Asp Gly Ile Asp Arg His Asp Arg Tyr  
 595 600 605  
 Ser Asn Arg Asn Val Val Phe Phe Asp Ile Asn Met Lys Gly Leu Asp  
 610 615 620  
 Gly Ile Gln Gly Pro Ile Tyr Val Gly Thr Gly Cys Val Phe Arg Arg  
 625 630 635 640  
 Gln Ala Leu Tyr Gly Phe Asp Ala Pro Lys Lys Lys Lys Pro Pro Gly  
 645 650 655  
 Lys Thr Cys Asn Cys Trp Pro Lys Trp Cys Cys Leu Cys Cys Gly Leu  
 660 665 670  
 Arg Lys Lys Ser Lys Thr Lys Ala Lys Asp Lys Lys Thr Asn Thr Lys  
 675 680 685  
 Glu Thr Ser Lys Gln Ile His Ala Leu Glu Asn Val Asp Glu Gly Val  
 690 695 700  
 Ile Val Pro Val Ser Asn Val Glu Lys Arg Ser Glu Ala Thr Gln Leu  
 705 710 715 720  
 Lys Leu Glu Lys Lys Phe Gly Gln Ser Pro Val Phe Val Ala Ser Ala  
 725 730 735  
 Val Leu Gln Asn Gly Gly Val Pro Arg Asn Ala Ser Pro Ala Cys Leu  
 740 745 750  
 Leu Arg Glu Ala Ile Gln Val Ile Ser Cys Gly Tyr Glu Asp Lys Thr  
 755 760 765  
 Glu Trp Gly Lys Glu Ile Gly Trp Ile Tyr Gly Ser Val Thr Glu Asp  
 770 775 780  
 Ile Leu Thr Gly Phe Lys Met His Cys His Gly Trp Arg Ser Val Tyr  
 785 790 795 800  
 Cys Met Pro Lys Arg Ala Ala Phe Lys Gly Ser Ala Pro Ile Asn Leu  
 805 810 815  
 Ser Asp Arg Leu His Gln Val Leu Arg Trp Ala Leu Gly Ser Val Glu  
 820 825 830  
 Ile Phe Leu Ser Arg His Cys Pro Ile Trp Tyr Gly Tyr Gly Gly Gly  
 835 840 845  
 Leu Lys Trp Leu Glu Arg Phe Ser Tyr Ile Asn Ser Val Val Tyr Pro  
 850 855 860  
 Trp Thr Ser Leu Pro Leu Ile Val Tyr Cys Ser Leu Pro Ala Val Cys  
 865 870 875 880  
 Leu Leu Thr Gly Lys Phe Ile Val Pro Glu Ile Ser Asn Tyr Ala Gly  
 885 890 895  
 Ile Leu Phe Met Leu Met Phe Ile Ser Ile Ala Val Thr Gly Ile Leu  
 900 905 910

Glu Met Gln Trp Gly Gly Val Gly Ile Asp Asp Trp Trp Arg Asn Glu  
 915 920 925  
 Gln Phe Trp Val Ile Gly Gly Ala Ser Ser His Leu Phe Ala Leu Phe  
 930 935 940  
 Gln Gly Leu Leu Lys Val Leu Ala Gly Val Asn Thr Asn Phe Thr Val  
 945 950 955 960  
 Thr Ser Lys Ala Ala Asp Asp Gly Ala Phe Ser Glu Leu Tyr Ile Phe  
 965 970 975  
 Lys Trp Thr Thr Leu Leu Ile Pro Pro Thr Thr Leu Leu Ile Ile Asn  
 980 985 990  
 Ile Ile Gly Val Ile Val Gly Val Ser Asp Ala Ile Ser Asn Gly Tyr  
 995 1000 1005  
 Asp Ser Trp Gly Pro Leu Phe Gly Arg Leu Phe Phe Ala Leu Trp Val  
 1010 1015 1020  
 Ile Val His Leu Tyr Pro Phe Leu Lys Gly Met Leu Gly Lys Gln Asp  
 1025 1030 1035 1040  
 Lys Met Pro Thr Ile Ile Val Val Trp Ser Ile Leu Leu Ala Ser Ile  
 1045 1050 1055  
 Leu Thr Leu Leu Trp Val Arg Val Asn Pro Phe Val Ala Lys Gly Gly  
 1060 1065 1070  
 Pro Val Leu Glu Ile Cys Gly Leu Asn Cys Gly Asn  
 1075 1080  
 <210> 25  
 <211> 685  
 <212> PRT  
 <213> Gossypium hirsutum  
 <400> 25  
 Arg Arg Trp Val Pro Phe Cys Lys Lys His Asn Val Glu Pro Arg Ala  
 1 5 10 15  
 Pro Glu Phe Tyr Phe Asn Glu Lys Ile Asp Tyr Leu Lys Asp Lys Val  
 20 25 30  
 His Pro Ser Phe Val Lys Glu Arg Arg Ala Met Lys Arg Glu Tyr Glu  
 35 40 45  
 Glu Phe Lys Val Arg Ile Asn Ala Leu Val Ala Lys Ala Gln Lys Lys  
 50 55 60  
 Pro Glu Glu Gly Trp Val Met Gln Asp Gly Thr Pro Trp Pro Gly Asn  
 65 70 75 80  
 Asn Thr Arg Asp His Pro Gly Met Ile Gln Val Tyr Leu Gly Ser Ala  
 85 90 95  
 Gly Ala Leu Asp Val Asp Gly Lys Glu Leu Pro Arg Leu Val Tyr Val  
 100 105 110

Ser Arg Glu Lys Arg Pro Gly Tyr Gln His His Lys Lys Ala Gly Ala  
 115 120 125  
 Glu Asn Ala Leu Val Arg Val Ser Ala Val Leu Thr Asn Ala Pro Phe  
 130 135 140  
 Ile Leu Asn Leu Asp Cys Asp His Tyr Ile Asn Asn Ser Lys Ala Met  
 145 150 155 160  
 Arg Glu Ala Met Cys Phe Leu Met Asp Pro Gln Phe Gly Lys Lys Leu  
 165 170 175  
 Cys Tyr Val Gln Phe Pro Gln Arg Phe Asp Gly Ile Asp Arg His Asp  
 180 185 190  
 Arg Tyr Ala Asn Arg Asn Val Val Phe Phe Asp Ile Asn Met Leu Gly  
 195 200 205  
 Leu Asp Gly Leu Gln Gly Pro Val Tyr Val Gly Thr Gly Cys Val Phe  
 210 215 220  
 Asn Arg Gln Ala Leu Tyr Gly Tyr Asp Pro Pro Val Ser Glu Lys Arg  
 225 230 235 240  
 Pro Lys Met Thr Cys Asp Cys Trp Pro Ser Trp Cys Cys Cys Cys Cys  
 245 250 255  
 Gly Gly Ser Arg Lys Lys Ser Lys Lys Lys Gly Glu Lys Lys Gly Leu  
 260 265 270  
 Leu Gly Gly Leu Leu Tyr Gly Lys Lys Lys Lys Met Met Gly Lys Asn  
 275 280 285  
 Tyr Val Lys Lys Gly Ser Ala Pro Val Phe Asp Leu Glu Glu Ile Glu  
 290 295 300  
 Glu Gly Leu Glu Gly Tyr Glu Glu Leu Glu Lys Ser Thr Leu Met Ser  
 305 310 315 320  
 Gln Lys Asn Phe Glu Lys Arg Phe Gly Gln Ser Pro Val Phe Ile Ala  
 325 330 335  
 Ser Thr Leu Met Glu Asn Gly Gly Leu Pro Glu Gly Thr Asn Ser Thr  
 340 345 350  
 Ser Leu Ile Lys Glu Ala Ile His Val Ile Ser Cys Gly Tyr Glu Glu  
 355 360 365  
 Lys Thr Glu Trp Gly Lys Glu Ile Gly Trp Ile Tyr Gly Ser Val Thr  
 370 375 380  
 Glu Asp Ile Leu Thr Gly Phe Lys Met His Cys Arg Gly Trp Lys Ser  
 385 390 395 400  
 Val Tyr Cys Val Pro Lys Arg Pro Ala Phe Lys Gly Ser Ala Pro Ile  
 405 410 415  
 Asn Leu Ser Asp Arg Leu His Gln Val Leu Arg Trp Ala Leu Gly Ser  
 420 425 430

Val Glu Ile Phe Leu Ser Arg His Cys Pro Leu Trp Tyr Gly Tyr Gly  
 435 440 445  
 Gly Lys Leu Lys Trp Leu Glu Arg Leu Ala Tyr Ile Asn Thr Ile Val  
 450 455 460  
 Tyr Pro Phe Thr Ser Ile Pro Leu Leu Ala Tyr Cys Thr Ile Pro Ala  
 465 470 475 480  
 Val Cys Leu Leu Thr Gly Lys Phe Ile Ile Pro Thr Leu Ser Asn Leu  
 485 490 495  
 Thr Ser Val Trp Phe Leu Ala Leu Phe Leu Ser Ile Ile Ala Thr Gly  
 500 505 510  
 Val Leu Glu Leu Arg Trp Ser Gly Val Ser Ile Gln Asp Trp Trp Arg  
 515 520 525  
 Asn Glu Gln Phe Trp Val Ile Gly Gly Val Ser Ala His Leu Phe Ala  
 530 535 540  
 Val Phe Gln Gly Leu Leu Lys Val Leu Ala Gly Val Asp Thr Asn Phe  
 545 550 555 560  
 Thr Val Thr Ala Lys Ala Ala Asp Asp Thr Glu Phe Gly Glu Leu Tyr  
 565 570 575  
 Leu Phe Lys Trp Thr Thr Leu Leu Ile Pro Pro Thr Thr Leu Ile Ile  
 580 585 590  
 Leu Asn Met Val Gly Val Val Ala Gly Val Ser Asp Ala Ile Asn Asn  
 595 600 605  
 Gly Tyr Gly Ser Trp Gly Pro Leu Phe Gly Lys Leu Phe Phe Ala Phe  
 610 615 620  
 Trp Val Ile Leu His Leu Tyr Pro Phe Leu Lys Gly Leu Met Gly Arg  
 625 630 635 640  
 Gln Asn Arg Thr Pro Thr Ile Val Val Leu Trp Ser Ile Leu Leu Ala  
 645 650 655  
 Ser Ile Phe Ser Leu Val Trp Val Arg Ile Asp Pro Phe Leu Pro Lys  
 660 665 670  
 Gln Thr Gly Pro Val Leu Lys Gln Cys Gly Val Glu Cys  
 675 680 685  
 <210> 26  
 <211> 1111  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 26  
 Met Ala Ser Thr Pro Pro Gln Thr Ser Lys Lys Val Arg Asn Asn Ser  
 1 5 10 15  
 Gly Ser Gly Gln Thr Val Lys Phe Ala Arg Arg Thr Ser Ser Gly Arg  
 20 25 30

Tyr Val Ser Leu Ser Arg Asp Asn Ile Glu Leu Ser Gly Glu Leu Ser  
                   35                                  40                                  45  
 Gly Asp Tyr Ser Asn Tyr Thr Val His Ile Pro Pro Thr Pro Asp Asn  
                   50                                  55                                  60  
 Gln Pro Met Ala Thr Lys Ala Glu Glu Gln Tyr Val Ser Asn Ser Leu  
                   65                                  70                                  75                                  80  
 Phe Thr Gly Gly Phe Asn Ser Val Thr Arg Ala His Leu Met Asp Lys  
                                   85                                  90                                  95  
 Val Ile Asp Ser Asp Val Thr His Pro Gln Met Ala Gly Ala Lys Gly  
                   100                                  105                                  110  
 Ser Ser Cys Ala Met Pro Ala Cys Asp Gly Asn Val Met Lys Asp Glu  
                   115                                  120                                  125  
 Arg Gly Lys Asp Val Met Pro Cys Glu Cys Arg Phe Lys Ile Cys Arg  
                   130                                  135                                  140  
 Asp Cys Phe Met Asp Ala Gln Lys Glu Thr Gly Leu Cys Pro Gly Cys  
                   145                                  150                                  155                                  160  
 Lys Glu Gln Tyr Lys Ile Gly Asp Leu Asp Asp Thr Pro Asp Tyr  
                                   165                                  170                                  175  
 Ser Ser Gly Ala Leu Pro Leu Pro Ala Pro Gly Lys Asp Gln Arg Gly  
                   180                                  185                                  190  
 Asn Asn Asn Asn Met Ser Met Met Lys Arg Asn Gln Asn Gly Glu Phe  
                   195                                  200                                  205  
 Asp His Asn Arg Trp Leu Phe Glu Thr Gln Gly Thr Tyr Gly Tyr Gly  
                   210                                  215                                  220  
 Asn Ala Tyr Trp Pro Gln Asp Glu Met Tyr Gly Asp Asp Met Asp Glu  
                   225                                  230                                  235                                  240  
 Gly Met Arg Gly Gly Met Val Glu Thr Ala Asp Lys Pro Trp Arg Pro  
                                   245                                  250                                  255  
 Leu Ser Arg Arg Ile Pro Ile Pro Ala Ala Ile Ile Ser Pro Tyr Arg  
                   260                                  265                                  270  
 Leu Leu Ile Val Ile Arg Phe Val Val Leu Cys Phe Phe Leu Thr Trp  
                   275                                  280                                  285  
 Arg Ile Arg Asn Pro Asn Glu Asp Ala Ile Trp Leu Trp Leu Met Ser  
                   290                                  295                                  300  
 Ile Ile Cys Glu Leu Trp Phe Gly Phe Ser Trp Ile Leu Asp Gln Ile  
                   305                                  310                                  315                                  320  
 Pro Lys Leu Cys Pro Ile Asn Arg Ser Thr Asp Leu Glu Val Leu Arg  
                                   325                                  330                                  335  
 Asp Lys Phe Asp Met Pro Ser Pro Ser Asn Pro Thr Gly Arg Ser Asp  
                   340                                  345                                  350

Leu Pro Gly Ile Asp Leu Phe Val Ser Thr Ala Asp Pro Glu Lys Glu  
 355 360 365  
 Pro Pro Leu Val Thr Ala Asn Thr Ile Leu Ser Ile Leu Ala Val Asp  
 370 375 380  
 Tyr Pro Val Glu Lys Val Ser Cys Tyr Leu Ser Asp Asp Gly Gly Ala  
 385 390 395 400  
 Leu Leu Ser Phe Glu Ala Met Ala Glu Ala Ala Ser Phe Ala Asp Leu  
 405 410 415  
 Trp Val Pro Phe Cys Arg Lys His Asn Ile Glu Pro Arg Asn Pro Asp  
 420 425 430  
 Ser Tyr Phe Ser Leu Lys Ile Asp Pro Thr Lys Asn Lys Ser Arg Ile  
 435 440 445  
 Asp Phe Val Lys Asp Arg Arg Lys Ile Lys Arg Glu Tyr Asp Glu Phe  
 450 455 460  
 Lys Val Arg Ile Asn Gly Leu Pro Asp Ser Ile Arg Arg Arg Ser Asp  
 465 470 475 480  
 Ala Phe Asn Ala Arg Glu Glu Met Lys Ala Leu Lys Gln Met Arg Glu  
 485 490 495  
 Ser Gly Gly Asp Pro Thr Glu Pro Val Lys Val Pro Lys Ala Thr Trp  
 500 505 510  
 Met Ala Asp Gly Thr His Trp Pro Gly Thr Trp Ala Ala Ser Thr Arg  
 515 520 525  
 Glu His Ser Lys Gly Asp His Ala Gly Ile Leu Gln Val Met Leu Lys  
 530 535 540  
 Pro Pro Ser Ser Asp Pro Leu Ile Gly Asn Ser Asp Asp Lys Val Ile  
 545 550 555 560  
 Asp Phe Ser Asp Thr Asp Thr Arg Leu Pro Met Phe Val Tyr Val Ser  
 565 570 575  
 Arg Glu Lys Arg Pro Gly Tyr Asp His Asn Lys Lys Ala Gly Ala Met  
 580 585 590  
 Asn Ala Leu Val Arg Ala Ser Ala Ile Leu Ser Asn Gly Pro Phe Ile  
 595 600 605  
 Leu Asn Leu Asp Cys Asp His Tyr Ile Tyr Asn Cys Lys Ala Val Arg  
 610 615 620  
 Glu Gly Met Cys Phe Met Met Asp Arg Gly Gly Glu Asp Ile Cys Tyr  
 625 630 635 640  
 Ile Gln Phe Pro Gln Arg Phe Glu Gly Ile Asp Pro Ser Asp Arg Tyr  
 645 650 655  
 Ala Asn Asn Asn Thr Val Phe Phe Asp Gly Asn Met Arg Ala Leu Asp  
 660 665 670

Gly Val Gln Gly Pro Val Tyr Val Gly Thr Gly Thr Met Phe Arg Arg  
 675 680 685  
 Phe Ala Leu Tyr Gly Phe Asp Pro Pro Asn Pro Asp Lys Leu Leu Glu  
 690 695 700  
 Lys Lys Glu Ser Glu Thr Glu Ala Leu Thr Thr Ser Asp Phe Asp Pro  
 705 710 715 720  
 Asp Leu Asp Val Thr Gln Leu Pro Lys Arg Phe Gly Asn Ser Thr Leu  
 725 730 735  
 Leu Ala Glu Ser Ile Pro Ile Ala Glu Phe Gln Gly Arg Pro Leu Ala  
 740 745 750  
 Asp His Pro Ala Val Lys Tyr Gly Arg Pro Pro Gly Ala Leu Arg Val  
 755 760 765  
 Pro Arg Asp Pro Leu Asp Ala Thr Thr Val Ala Glu Ser Val Ser Val  
 770 775 780  
 Ile Ser Cys Trp Tyr Glu Asp Lys Thr Glu Trp Gly Asp Arg Val Gly  
 785 790 795 800  
 Trp Ile Tyr Gly Ser Val Thr Glu Asp Val Val Thr Gly Tyr Arg Met  
 805 810 815  
 His Asn Arg Gly Trp Arg Ser Val Tyr Cys Ile Thr Lys Arg Asp Ser  
 820 825 830  
 Phe Arg Gly Ser Ala Pro Ile Asn Leu Thr Asp Arg Leu His Gln Val  
 835 840 845  
 Leu Arg Trp Ala Thr Gly Ser Val Glu Ile Phe Phe Ser Arg Asn Asn  
 850 855 860  
 Ala Ile Leu Ala Ser Lys Arg Leu Lys Phe Leu Gln Arg Leu Ala Tyr  
 865 870 875 880  
 Leu Asn Val Gly Ile Tyr Pro Phe Thr Ser Leu Phe Leu Ile Leu Tyr  
 885 890 895  
 Cys Phe Leu Pro Ala Phe Ser Leu Phe Ser Gly Gln Phe Ile Val Arg  
 900 905 910  
 Thr Leu Ser Ile Ser Phe Leu Val Tyr Leu Leu Met Ile Thr Ile Cys  
 915 920 925  
 Leu Ile Gly Leu Ala Val Leu Glu Val Lys Trp Ser Gly Ile Gly Leu  
 930 935 940  
 Glu Glu Trp Trp Arg Asn Glu Gln Trp Trp Leu Ile Ser Gly Thr Ser  
 945 950 955 960  
 Ser His Leu Tyr Ala Val Val Gln Gly Val Leu Lys Val Ile Ala Gly  
 965 970 975  
 Ile Glu Ile Ser Phe Thr Leu Thr Thr Lys Ser Gly Gly Asp Asp Asn  
 980 985 990



Glu Asp Ile Tyr Ala Asp Leu Tyr Ile Val Lys Trp Ser Ser Leu Met  
 995 1000 1005  
 Ile Pro Pro Ile Val Ile Ala Met Val Asn Ile Ile Ala Ile Val Val  
 1010 1015 1020  
 Ala Phe Ile Arg Thr Ile Tyr Gln Ala Val Pro Gln Trp Ser Lys Leu  
 1025 1030 1035 1040  
 Ile Gly Gly Ala Phe Phe Ser Phe Trp Val Leu Ala His Leu Tyr Pro  
 1045 1050 1055  
 Phe Ala Lys Gly Leu Met Gly Arg Arg Gly Lys Thr Pro Thr Ile Val  
 1060 1065 1070  
 Phe Val Trp Ala Gly Leu Ile Ala Ile Thr Ile Ser Leu Leu Trp Thr  
 1075 1080 1085  
 Ala Ile Asn Pro Asn Thr Gly Pro Ala Ala Ala Ala Glu Gly Val Gly  
 1090 1095 1100  
 Gly Gly Gly Phe Gln Phe Pro  
 1105 1110  
 <210> 27  
 <211> 1026  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 27  
 Met Glu Ala Ser Ala Gly Leu Val Ala Gly Ser His Asn Arg Asn Glu  
 1 5 10 15  
 Leu Val Val Ile His Asn His Glu Glu Pro Lys Pro Leu Lys Asn Leu  
 20 25 30  
 Asp Gly Gln Phe Cys Glu Ile Cys Gly Asp Gln Ile Gly Leu Thr Val  
 35 40 45  
 Glu Gly Asp Leu Phe Val Ala Cys Asn Glu Cys Gly Phe Pro Ala Cys  
 50 55 60  
 Arg Pro Cys Tyr Glu Tyr Glu Arg Arg Glu Gly Thr Gln Asn Cys Pro  
 65 70 75 80  
 Gln Cys Lys Thr Arg Tyr Lys Arg Leu Arg Gly Ser Pro Arg Val Glu  
 85 90 95  
 Gly Asp Glu Asp Glu Glu Asp Ile Asp Asp Ile Glu Tyr Glu Phe Asn  
 100 105 110  
 Ile Glu His Glu Gln Asp Lys His Lys His Ser Ala Glu Ala Met Leu  
 115 120 125  
 Tyr Gly Lys Met Ser Tyr Gly Arg Gly Pro Glu Asp Asp Glu Asn Gly  
 130 135 140  
 Arg Phe Pro Pro Val Ile Ala Gly Gly His Ser Gly Glu Phe Pro Val  
 145 150 155 160

Gly Gly Gly Tyr Gly Asn Gly Glu His Gly Leu His Lys Arg Val His  
 165 170 175  
 Pro Tyr Pro Ser Ser Glu Ala Gly Ser Glu Gly Gly Trp Arg Glu Arg  
 180 185 190  
 Met Asp Asp Trp Lys Leu Gln His Gly Asn Leu Gly Pro Glu Pro Asp  
 195 200 205  
 Asp Asp Pro Glu Met Gly Leu Ile Asp Glu Ala Arg Gln Pro Leu Ser  
 210 215 220  
 Arg Lys Val Pro Ile Ala Ser Ser Lys Ile Asn Pro Tyr Arg Met Val  
 225 230 235 240  
 Ile Val Ala Arg Leu Val Ile Leu Ala Val Phe Leu Arg Tyr Arg Leu  
 245 250 255  
 Leu Asn Pro Val His Asp Ala Leu Gly Leu Trp Leu Thr Ser Val Ile  
 260 265 270  
 Cys Glu Ile Trp Phe Ala Val Ser Trp Ile Leu Asp Gln Phe Pro Lys  
 275 280 285  
 Trp Phe Pro Ile Glu Arg Glu Thr Tyr Leu Asp Arg Leu Ser Leu Arg  
 290 295 300  
 Tyr Glu Arg Glu Gly Glu Pro Asn Met Leu Ala Pro Val Asp Val Phe  
 305 310 315 320  
 Val Ser Thr Val Asp Pro Leu Lys Glu Pro Pro Leu Val Thr Ser Asn  
 325 330 335  
 Thr Val Leu Ser Ile Leu Ala Met Asp Tyr Pro Val Glu Lys Ile Ser  
 340 345 350  
 Cys Tyr Val Ser Asp Asp Gly Ala Ser Met Leu Thr Phe Glu Ser Leu  
 355 360 365  
 Ser Glu Thr Ala Glu Phe Ala Arg Lys Trp Val Pro Phe Cys Lys Lys  
 370 375 380  
 Phe Ser Ile Glu Pro Arg Ala Pro Glu Met Tyr Phe Thr Leu Lys Val  
 385 390 395 400  
 Asp Tyr Leu Gln Asp Lys Val His Pro Thr Phe Val Lys Glu Arg Arg  
 405 410 415  
 Ala Met Lys Arg Glu Tyr Glu Glu Phe Lys Val Arg Ile Asn Ala Gln  
 420 425 430  
 Val Ala Lys Ala Ser Lys Val Pro Leu Glu Gly Trp Ile Met Gln Asp  
 435 440 445  
 Gly Thr Pro Trp Pro Gly Asn Asn Thr Lys Asp His Pro Gly Met Ile  
 450 455 460  
 Gln Val Phe Leu Gly His Ser Gly Gly Phe Asp Val Glu Gly His Glu  
 465 470 475 480

Leu Pro Arg Leu Val Tyr Val Ser Arg Glu Lys Arg Pro Gly Phe Gln  
 485 490 495  
 His His Lys Lys Ala Gly Ala Met Asn Ala Leu Val Arg Val Ala Gly  
 500 505 510  
 Val Leu Thr Asn Ala Pro Phe Met Leu Asn Leu Asp Cys Asp His Tyr  
 515 520 525  
 Val Asn Asn Ser Lys Ala Val Arg Glu Ala Met Cys Phe Leu Met Asp  
 530 535 540  
 Pro Gln Ile Gly Lys Lys Val Cys Tyr Val Gln Phe Pro Gln Arg Phe  
 545 550 555 560  
 Asp Gly Ile Asp Thr Asn Asp Arg Tyr Ala Asn Arg Asn Thr Val Phe  
 565 570 575  
 Phe Asp Ile Asn Met Lys Gly Leu Asp Gly Ile Gln Gly Pro Val Tyr  
 580 585 590  
 Val Gly Thr Gly Cys Val Phe Lys Arg Gln Ala Leu Tyr Gly Tyr Glu  
 595 600 605  
 Pro Pro Lys Gly Pro Lys Arg Pro Lys Met Ile Ser Cys Gly Cys Cys  
 610 615 620  
 Pro Cys Phe Gly Arg Arg Arg Lys Asn Lys Lys Phe Ser Lys Asn Asp  
 625 630 635 640  
 Met Asn Gly Asp Val Ala Ala Leu Gly Gly Ala Glu Gly Asp Lys Glu  
 645 650 655  
 His Leu Met Phe Glu Met Asn Phe Glu Lys Thr Phe Gly Gln Ser Ser  
 660 665 670  
 Ile Phe Val Thr Ser Thr Leu Met Glu Glu Gly Gly Val Pro Pro Ser  
 675 680 685  
 Ser Ser Pro Ala Val Leu Leu Lys Glu Ala Ile His Val Ile Ser Cys  
 690 695 700  
 Gly Tyr Glu Asp Lys Thr Glu Trp Gly Thr Glu Leu Gly Trp Ile Tyr  
 705 710 715 720  
 Gly Ser Ile Thr Glu Asp Ile Leu Thr Gly Phe Lys Met His Cys Arg  
 725 730 735  
 Gly Trp Arg Ser Ile Tyr Cys Met Pro Lys Arg Pro Ala Phe Lys Gly  
 740 745 750  
 Ser Ala Pro Ile Asn Leu Ser Asp Arg Leu Asn Gln Val Leu Arg Trp  
 755 760 765  
 Ala Leu Gly Ser Val Glu Ile Phe Phe Ser Arg His Ser Pro Leu Trp  
 770 775 780  
 Tyr Gly Tyr Lys Gly Gly Lys Leu Lys Trp Leu Glu Arg Phe Ala Tyr  
 785 790 795 800

Ala Asn Thr Thr Ile Tyr Pro Phe Thr Ser Ile Pro Leu Leu Ala Tyr  
                     805                    810                    815  
 Cys Ile Leu Pro Ala Ile Cys Leu Leu Thr Asp Lys Phe Ile Met Pro  
                     820                    825                    830  
 Pro Ile Ser Thr Phe Ala Ser Leu Phe Phe Ile Ser Leu Phe Met Ser  
                     835                    840                    845  
 Ile Ile Val Thr Gly Ile Leu Glu Leu Arg Trp Ser Gly Val Ser Ile  
                     850                    855                    860  
 Glu Glu Trp Trp Arg Asn Glu Gln Phe Trp Val Ile Gly Gly Ile Ser  
                     865                    870                    875                    880  
 Ala His Leu Phe Ala Val Val Gln Gly Leu Leu Lys Ile Leu Ala Gly  
                     885                    890                    895  
 Ile Asp Thr Asn Phe Thr Val Thr Ser Lys Ala Thr Asp Asp Asp  
                     900                    905                    910  
 Phe Gly Glu Leu Tyr Ala Phe Lys Trp Thr Thr Leu Leu Ile Pro Pro  
                     915                    920                    925  
 Thr Thr Val Leu Ile Ile Asn Ile Val Gly Val Val Ala Gly Ile Ser  
                     930                    935                    940  
 Asp Ala Ile Asn Asn Gly Tyr Gln Ser Trp Gly Pro Leu Phe Gly Lys  
                     945                    950                    955                    960  
 Leu Phe Phe Ser Phe Trp Val Ile Val His Leu Tyr Pro Phe Leu Lys  
                     965                    970                    975  
 Gly Leu Met Gly Arg Gln Asn Arg Thr Pro Thr Ile Val Val Ile Trp  
                     980                    985                    990  
 Ser Val Leu Leu Ala Ser Ile Phe Ser Leu Leu Trp Val Arg Ile Asp  
                     995                    1000                    1005  
 Pro Phe Val Leu Lys Thr Lys Gly Pro Asp Thr Ser Lys Cys Gly Ile  
                     1010                    1015                    1020  
 Asn Cys  
 1025

<210> 28  
 <211> 701  
 <212> PRT  
 <213> *Gossypium hirsutum*

<400> 28  
 Asp Tyr Pro Val Glu Lys Val Ser Cys Tyr Val Ser Asp Asp Gly Ala  
   1                    5                    10                    15  
 Ala Met Leu Thr Phe Glu Ala Leu Ser Glu Thr Ser Glu Phe Ala Arg  
                     20                    25                    30  
 Lys Trp Val Pro Phe Cys Lys Lys Tyr Asn Ile Glu Pro Arg Ala Pro  
                     35                    40                    45

Glu Trp Tyr Phe Ala Gln Lys Ile Asp Tyr Leu Lys Asp Lys Val Gln  
 50 55 60  
 Thr Ser Phe Val Lys Glu Arg Arg Ala Met Lys Arg Glu Tyr Glu Glu  
 65 70 75 80  
 Phe Lys Val Arg Val Asn Gly Leu Val Ala Lys Ala Gln Lys Val Pro  
 85 90 95  
 Glu Glu Gly Trp Ile Met Gln Asp Gly Thr Pro Trp Pro Gly Asn Asn  
 100 105 110  
 Thr Arg Asp His Pro Gly Met Ile Gln Val Phe Leu Gly Gln Ser Gly  
 115 120 125  
 Gly Leu Asp Ala Glu Gly Asn Glu Leu Pro Arg Leu Val Tyr Val Ser  
 130 135 140  
 Arg Glu Lys Arg Pro Gly Phe Gln His His Lys Lys Ala Gly Ala Met  
 145 150 155 160  
 Asn Ala Leu Val Arg Val Ser Ala Val Leu Thr Asn Gly Ala Phe Leu  
 165 170 175  
 Leu Asn Leu Asp Cys Asp His Tyr Ile Asn Asn Ser Lys Ala Leu Arg  
 180 185 190  
 Glu Ala Met Cys Phe Leu Met Asp Pro Asn Leu Gly Lys Gln Val Cys  
 195 200 205  
 Tyr Val Gln Phe Pro Gln Arg Phe Asp Gly Ile Asp Arg Asn Asp Arg  
 210 215 220  
 Tyr Ala Asn Arg Asn Thr Val Phe Phe Asp Ile Asn Leu Arg Gly Leu  
 225 230 235 240  
 Asp Gly Ile Gln Gly Pro Val Tyr Val Gly Thr Gly Cys Val Phe Asn  
 245 250 255  
 Arg Thr Ala Leu Tyr Gly Tyr Glu Pro Pro Leu Lys Pro Lys His Arg  
 260 265 270  
 Lys Thr Gly Ile Leu Ser Ser Leu Cys Gly Gly Ser Arg Lys Lys Ser  
 275 280 285  
 Ser Lys Ser Ser Lys Lys Gly Ser Asp Lys Lys Lys Ser Gly Lys His  
 290 295 300  
 Val Asp Ser Thr Val Pro Val Phe Asn Leu Glu Asp Ile Glu Glu Gly  
 305 310 315 320  
 Val Glu Gly Ala Gly Phe Asp Asp Glu Lys Ser Leu Leu Met Ser Gln  
 325 330 335  
 Met Ser Leu Glu Lys Arg Phe Gly Gln Ser Ala Val Phe Val Ala Ser  
 340 345 350  
 Thr Leu Met Glu Asn Gly Gly Val Pro Gln Ser Ala Thr Pro Glu Thr  
 355 360 365

Leu Leu Lys Glu Ala Ile His Val Ile Ser Cys Gly Tyr Glu Asp Lys  
 370 375 380  
 Thr Asp Trp Gly Ser Glu Ile Gly Trp Ile Tyr Gly Ser Val Thr Glu  
 385 390 395 400  
 Asp Ile Leu Thr Gly Phe Lys Met His Ala Arg Gly Trp Arg Ser Ile  
 405 410 415  
 Tyr Cys Met Pro Lys Arg Pro Ala Phe Lys Gly Ser Ala Pro Ile Asn  
 420 425 430  
 Leu Ser Asp Arg Leu Asn Gln Val Leu Arg Trp Ala Leu Gly Ser Val  
 435 440 445  
 Glu Ile Leu Phe Ser Arg His Cys Pro Ile Trp Tyr Gly Tyr Ser Gly  
 450 455 460  
 Arg Leu Lys Trp Leu Glu Arg Phe Ala Tyr Val Asn Thr Thr Ile Tyr  
 465 470 475 480  
 Pro Val Thr Ala Ile Pro Leu Leu Met Tyr Cys Thr Leu Pro Ala Val  
 485 490 495  
 Cys Leu Leu Thr Asn Lys Phe Ile Ile Pro Gln Ile Ser Asn Leu Ala  
 500 505 510  
 Ser Ile Trp Phe Ile Ser Leu Phe Leu Ser Ile Phe Ala Thr Gly Ile  
 515 520 525  
 Leu Lys Met Lys Trp Asn Gly Val Gly Ile Asp Gln Trp Trp Arg Asn  
 530 535 540  
 Glu Gln Phe Trp Val Ile Gly Gly Val Ser Ala His Leu Phe Ala Val  
 545 550 555 560  
 Phe Gln Gly Leu Leu Lys Val Leu Ala Gly Ile Asp Thr Asn Phe Thr  
 565 570 575  
 Val Thr Ser Lys Ala Ser Asp Glu Asp Gly Asp Phe Ala Glu Leu Tyr  
 580 585 590  
 Met Phe Lys Trp Thr Thr Leu Leu Ile Pro Pro Thr Thr Leu Leu Ile  
 595 600 605  
 Ile Asn Leu Val Gly Val Val Ala Gly Ile Ser Tyr Val Ile Asn Ser  
 610 615 620  
 Gly Tyr Gln Ser Trp Gly Pro Leu Phe Gly Lys Leu Phe Phe Ala Phe  
 625 630 635 640  
 Trp Val Ile Ile His Leu Tyr Pro Phe Leu Lys Gly Leu Met Gly Arg  
 645 650 655  
 Gln Asn Arg Thr Pro Thr Ile Val Val Val Trp Ser Ile Leu Leu Ala  
 660 665 670  
 Ser Ile Phe Ser Leu Leu Trp Val Arg Ile Asp Pro Phe Thr Thr Arg  
 675 680 685

Val Thr Gly Pro Asp Val Glu Gln Cys Gly Ile Asn Cys  
 690 695 700

<210> 29  
 <211> 1081  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 29  
 Arg Pro Arg Leu Ile Ala Gly Ser His Asn Arg Asn Glu Phe Val Leu  
 1 5 10 15

Ile Asn Ala Asp Glu Asn Ala Arg Ile Arg Ser Val Gln Glu Leu Ser  
 20 25 30

Gly Gln Thr Cys Gln Ile Cys Arg Asp Glu Ile Glu Leu Thr Val Asp  
 35 40 45

Gly Glu Pro Phe Val Ala Cys Asn Glu Cys Ala Phe Pro Val Cys Arg  
 50 55 60

Pro Cys Tyr Glu Tyr Glu Arg Arg Glu Gly Asn Gln Ala Cys Pro Gln  
 65 70 75 80

Cys Lys Thr Arg Phe Lys Arg Leu Lys Gly Ser Pro Arg Val Glu Gly  
 85 90 95

Asp Glu Glu Glu Asp Asp Ile Asp Asp Leu Asp Asn Glu Phe Glu Tyr  
 100 105 110

Gly Asn Asn Gly Ile Gly Phe Asp Gln Val Ser Glu Gly Met Ser Ile  
 115 120 125

Ser Arg Arg Asn Ser Gly Phe Pro Gln Ser Asp Leu Asp Ser Ala Pro  
 130 135 140

Pro Gly Ser Gln Ile Pro Leu Leu Thr Tyr Gly Asp Glu Asp Val Glu  
 145 150 155 160

Ile Ser Ser Asp Arg His Ala Leu Ile Val Pro Pro Ser Leu Gly Gly  
 165 170 175

His Gly Asn Arg Val His Pro Val Ser Leu Ser Asp Pro Thr Val Ala  
 180 185 190

Ala His Arg Arg Leu Met Val Pro Gln Lys Asp Leu Ala Val Tyr Gly  
 195 200 205

Tyr Gly Ser Val Ala Trp Lys Asp Arg Met Glu Glu Trp Lys Arg Lys  
 210 215 220

Gln Asn Glu Lys Leu Gln Val Val Arg His Glu Gly Asp Pro Asp Phe  
 225 230 235 240

Glu Asp Gly Asp Asp Ala Asp Phe Pro Met Met Asp Glu Gly Arg Gln  
 245 250 255

Pro Leu Ser Met Lys Ile Pro Ile Lys Ser Ser Lys Ile Asn Pro Tyr  
 260 265 270

Arg Met Leu Ile Val Leu Arg Leu Val Ile Leu Gly Leu Phe Phe His  
 275 280 285  
 Tyr Arg Ile Leu His Pro Val Lys Asp Ala Tyr Ala Leu Trp Leu Ile  
 290 295 300  
 Ser Val Ile Cys Glu Ile Trp Phe Ala Val Ser Trp Val Leu Asp Gln  
 305 310 315 320  
 Phe Pro Lys Trp Tyr Pro Ile Glu Arg Glu Thr Tyr Leu Asp Arg Leu  
 325 330 335  
 Ser Leu Arg Tyr Glu Lys Glu Gly Lys Pro Ser Gly Leu Ser Pro Val  
 340 345 350  
 Asp Val Phe Val Ser Thr Val Asp Pro Leu Lys Glu Pro Pro Leu Ile  
 355 360 365  
 Thr Ala Asn Thr Val Leu Ser Ile Leu Ala Val Asp Tyr Pro Val Asp  
 370 375 380  
 Lys Val Ala Cys Tyr Val Ser Asp Asp Gly Ala Ala Met Leu Thr Phe  
 385 390 395 400  
 Glu Ala Leu Ser Glu Thr Ala Glu Phe Ala Arg Lys Trp Val Pro Phe  
 405 410 415  
 Cys Lys Lys Tyr Cys Ile Glu Pro Arg Ala Pro Glu Trp Tyr Phe Cys  
 420 425 430  
 His Lys Met Asp Tyr Leu Lys Asn Lys Val His Pro Ala Phe Val Arg  
 435 440 445  
 Glu Arg Arg Ala Met Lys Arg Asp Tyr Glu Glu Phe Lys Val Lys Ile  
 450 455 460  
 Asn Ala Leu Val Ala Thr Ala Gln Lys Val Pro Glu Asp Gly Trp Thr  
 465 470 475 480  
 Met Gln Asp Gly Thr Pro Trp Pro Gly Asn Ser Val Arg Asp His Pro  
 485 490 495  
 Gly Met Ile Gln Val Phe Leu Gly Ser Asp Gly Val Arg Asp Val Glu  
 500 505 510  
 Asn Asn Glu Leu Pro Arg Leu Val Tyr Val Ser Arg Glu Lys Arg Pro  
 515 520 525  
 Gly Phe Asp His His Lys Lys Ala Gly Ala Met Asn Ser Leu Ile Arg  
 530 535 540  
 Val Ser Gly Val Leu Ser Asn Ala Pro Tyr Leu Leu Asn Val Asp Cys  
 545 550 555 560  
 Asp His Tyr Ile Asn Asn Ser Lys Ala Leu Arg Glu Ala Met Cys Phe  
 565 570 575  
 Met Met Asp Pro Gln Ser Gly Lys Lys Ile Cys Tyr Val Gln Phe Pro  
 580 585 590



Gln Arg Phe Asp Gly Ile Asp Arg His Asp Arg Tyr Ser Asn Arg Asn  
 595 600 605  
 Val Val Phe Phe Asp Ile Asn Met Lys Gly Leu Asp Gly Leu Gln Gly  
 610 615 620  
 Pro Ile Tyr Val Gly Thr Gly Cys Val Phe Arg Arg Gln Ala Leu Tyr  
 625 630 635 640  
 Gly Phe Asp Ala Pro Lys Lys Lys Lys Gly Pro Arg Lys Thr Cys Asn  
 645 650 655  
 Cys Trp Pro Lys Trp Cys Leu Leu Cys Phe Gly Ser Arg Lys Asn Arg  
 660 665 670  
 Lys Ala Lys Thr Val Ala Ala Asp Lys Lys Lys Lys Asn Arg Glu Ala  
 675 680 685  
 Ser Lys Gln Ile His Ala Leu Glu Asn Ile Glu Glu Gly Arg Gly His  
 690 695 700  
 Lys Val Leu Asn Val Glu Gln Ser Thr Glu Ala Met Gln Met Lys Leu  
 705 710 715 720  
 Gln Lys Lys Tyr Gly Gln Ser Pro Val Phe Val Ala Ser Ala Arg Leu  
 725 730 735  
 Glu Asn Gly Gly Met Ala Arg Asn Ala Ser Pro Ala Cys Leu Leu Lys  
 740 745 750  
 Glu Ala Ile Gln Val Ile Ser Arg Gly Tyr Glu Asp Lys Thr Glu Trp  
 755 760 765  
 Gly Lys Glu Ile Gly Trp Ile Tyr Gly Ser Val Thr Glu Asp Ile Leu  
 770 775 780  
 Thr Gly Ser Lys Met His Ser His Gly Trp Arg His Val Tyr Cys Thr  
 785 790 795 800  
 Pro Lys Leu Ala Ala Phe Lys Gly Ser Ala Pro Ile Asn Leu Ser Asp  
 805 810 815  
 Arg Leu His Gln Val Leu Arg Trp Ala Leu Gly Ser Val Glu Ile Phe  
 820 825 830  
 Leu Ser Arg His Cys Pro Ile Trp Tyr Gly Tyr Gly Gly Leu Lys  
 835 840 845  
 Trp Leu Glu Arg Leu Ser Tyr Ile Asn Ser Val Val Tyr Pro Trp Thr  
 850 855 860  
 Ser Leu Pro Leu Ile Val Tyr Cys Ser Leu Pro Ala Ile Cys Leu Leu  
 865 870 875 880  
 Thr Gly Lys Phe Ile Val Pro Glu Ile Ser Asn Tyr Ala Ser Ile Leu  
 885 890 895  
 Phe Met Ala Leu Phe Ser Ser Ile Ala Ile Thr Gly Ile Leu Glu Met  
 900 905 910

Gln Trp Gly Lys Val Gly Ile Asp Asp Trp Trp Arg Asn Glu Gln Phe  
 915 920 925  
 Trp Val Ile Gly Gly Val Ser Ala His Leu Phe Ala Leu Phe Gln Gly  
 930 935 940  
 Leu Leu Lys Val Leu Ala Gly Val Asp Thr Asn Phe Thr Val Thr Ser  
 945 950 955 960  
 Lys Ala Ala Asp Asp Gly Glu Phe Ser Asp Leu Tyr Leu Phe Lys Trp  
 965 970 975  
 Thr Ser Leu Leu Ile Pro Pro Met Thr Leu Leu Ile Ile Asn Val Ile  
 980 985 990  
 Gly Val Ile Val Gly Val Ser Asp Ala Ile Ser Asn Gly Tyr Asp Ser  
 995 1000 1005  
 Trp Gly Pro Leu Phe Gly Arg Leu Phe Phe Ala Leu Trp Val Ile Ile  
 1010 1015 1020  
 His Leu Tyr Pro Phe Leu Lys Gly Leu Leu Gly Lys Gln Asp Arg Met  
 1025 1030 1035 1040  
 Pro Thr Ile Ile Val Val Trp Ser Ile Leu Leu Ala Ser Ile Leu Thr  
 1045 1050 1055  
 Leu Leu Trp Val Arg Val Asn Pro Phe Val Ala Lys Gly Gly Pro Ile  
 1060 1065 1070  
 Leu Glu Ile Cys Gly Leu Asp Cys Leu  
 1075 1080

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

[illegible]

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

# INTERNATIONAL SEARCH REPORT

International Application No  
PC., US 99/15871

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC 7    C12N15/54    C12N1/21    C12N9/10    C12Q1/48    C12Q1/68		
According to International Patent Classification (IPC) or to both national classification and IPO		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) IPC 7    C12N    C12Q		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 00549 A (WILLIAMSON RICHARD EDWARD ;PENG LIANGCAI (AU); ARIOLI ANTONIO (AU)) 8 January 1998 (1998-01-08) see SEQ ID NOs:1-12 <div style="text-align: center; margin-top: 10px;">           ---            -/--         </div>	1,2,4-7, 10-17
<div style="display: flex; justify-content: space-between;"> <span><input checked="" type="checkbox"/> Further documents are listed in the continuation of box C.</span> <span><input checked="" type="checkbox"/> Patent family members are listed in annex.</span> </div>		
<div style="display: flex;"> <div style="flex: 1;"> <p>* Special categories of cited documents :</p> <p>*A* document defining the general state of the art which is not considered to be of particular relevance</p> <p>*E* earlier document but published on or after the international filing date</p> <p>*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>*O* document referring to an oral disclosure, use, exhibition or other means</p> <p>*P* document published prior to the international filing date but later than the priority date claimed</p> </div> <div style="flex: 1;"> <p>*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>*Z* document member of the same patent family</p> </div> </div>		
Date of the actual completion of the international search  <div style="text-align: center; font-size: 1.2em;">9 February 2000</div>		Date of mailing of the international search report  <div style="text-align: center; font-size: 1.2em;">23.02.00</div>
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3018		Authorized officer  <div style="text-align: center; font-size: 1.2em;">Maddox, A</div>

## INTERNATIONAL SEARCH REPORT

International Application No

PL., US 99/15871

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ARIOLI, T., ET AL.: "Arabidopsis thaliana cellulose synthase catalytic subunit (RSW1) gene complete cds" EMBL ACCESSION NO:AF027172, 3 February 1998 (1998-02-03), XP002124282 the whole document	1,4-6, 13-17
X	-& ARIOLI, T. ET AL.: "Molecular analysis of cellulose biosynthesis in Arabidopsis" SCIENCE, vol. 279, 30 January 1998 (1998-01-30), pages 717-720, XP002124283 the whole document & ARIOLI, T., ET AL.: "Cellulose synthase catalytic subunit" TREMBL ACCESSION NO:048946, 1 June 1998 (1998-06-01), ---	6,13-17
X	ARIOLI, T., ET AL.: "Arabidopsis thaliana cellulose synthase catalytic subunit (Ath-A) mRNA, complete cds" EMBL ACCESSION NO: AF027173, 3 February 1998 (1998-02-03), XP002129994 the whole document	7,10-17
X	-& ARIOLI, T., ET AL.: "Molecular analysis of cellulose biosynthesis in Arabidopsis" SCIENCE, vol. 279, 30 January 1998 (1998-01-30), pages 717-720, XP002124283 the whole document & ARIOLI, T., ET AL.: "CELLULOSE SYNTHASE CATALYTIC SUBUNIT." TREMBL ACCESSION NO.048497, 1 June 1998 (1998-06-01), ---	12-17
X	ARIOLI, T., ET AL.: "Arabidopsis thaliana cellulose synthase catalytic subunit (Ath-B) mRNA, complete cds" EMBL ACCESSION NO:AF027174, 3 February 1998 (1998-02-03), XP002124284 the whole document	1,2,4,5, 13-17
X	-& ARIOLI T., ET AL.: "Molecular analysis of cellulose biosynthesis in Arabidopsis " SCIENCE, vol. 279, 30 January 1998 (1998-01-30), pages 717-720, XP002124283 the whole document & ARIOLI, T., ET AL.: "Cellulose synthase catalytic subunit" TREMBL ACCESSION NO:048948, 1 June 1998 (1998-06-01), ---	6,13-17

-/--

## INTERNATIONAL SEARCH REPORT

International Application No

PC, US 99/15871

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 18949 A (CALGENE INC ;PEAR JULIE R (US); STALKER DAVID M (US); DELMER DEBOR) 7 May 1998 (1998-05-07) figures 3A-C,6A-E,7A-D ---	1,2,4-7, 10-17
X	PEAR, J.R., ET AL.: "Gossypium hirsutum cellulose synthase (celA2) mRNA, partial cds" EMBL ACCESSION NO:U58284, 13 December 1996 (1996-12-13), XP002124438	1,2,4-7, 10-17
X	-& PEAR, J.R., ET AL.: "HIGHER PLANTS CONTAIN HOMOLOGS OF THE BACTERIAL CELA GENES ENCODINGTHE CATALYTIC SUBUNIT OF CELLULOSE SYNTHASE" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 93, October 1996 (1996-10), pages 12637-12642, XP002061424	6,12-17
Y	see whole document particularly footnote left col page 12639  & PEAR, J.R., ET AL.: "Cellulose synthase fragment" TREMBL ACCESSION NO:P93156, 1 May 1997 (1997-05-01), ---	1,2,4-6
X	PEAR, J.R., ET AL.: "Gossypium hirsutum cellulose synthase (celA1) mRNA, complete cds" EMBL ACCESSION NO:U58283, 13 December 1996 (1996-12-13), XP002124439 the whole document	1,2,4-6, 13-17
X	-& PEAR, J.R., ET AL.: "HIGHER PLANTS CONTAIN HOMOLOGS OF THE BACTERIAL CELA GENES ENCODINGTHE CATALYTIC SUBUNIT OF CELLULOSE SYNTHASE" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 93, October 1996 (1996-10), pages 12637-12642, XP002061424 the whole document & PEAR, J.R. ET AL.: "Cellulose synthase" TREMBL ACCESSION NO:P93155, 1 May 1997 (1997-05-01), ---	6,13-17

-/--

## INTERNATIONAL SEARCH REPORT

International Application No

PL, /US 99/15871

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>WU, L., ET AL.: "Arabidopsis thaliana cellulose synthase mRNA, partial cds" EMBL ACCESSION NO: AF062485, 18 May 1998 (1998-05-18), XP002129995 the whole document</p> <p>-&amp; WU, L., ET AL.: "AraxCelA, a new member of cellulose synthase gene family from Arabidopsis thaliana (accession no. AF062485) (PGR 98-113)" PLANT PHYSIOLOGY 117:1125, July 1998 (1998-07), XP002130048</p> <p>---</p>	7,10-17
X	<p>DBEST DATABASE ID:37681, 2 December 1993 (1993-12-02), XP002124440 the whole document</p> <p>&amp; SASAKI, T.: "Rice cDNA, partial sequence (R1814-1A)" EMBL ACCESSION NO:D24381, 29 November 1993 (1993-11-29),</p> <p>---</p>	1,2,4-6, 16,17
X	<p>DATABASE DBEST ID:1473188, 20 January 1998 (1998-01-20), XP002129996 &amp; NAHM, B.H., ET AL.: "96AS0237 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa cDNA clone 96AS0237." EMBL ACCESSION NO:AA751514, 21 January 1998 (1998-01-21),</p> <p>---</p>	1,2,4-6, 16,17
X	<p>SASAKI, T., ET AL.: DATABASE DBEST ID:75334, 15 November 1994 (1994-11-15), XP002129997</p> <p>&amp; SASAKI, T. ET AL.: "Rice cDNA, partial sequence (S3630_1A)." EMBL ACCESSION NO:D41261, 13 November 1994 (1994-11-13),</p> <p>---</p>	7,10-12, 16,17
X	<p>SASAKI, T., ET AL.: DATABASE DBEST ID:75839, 15 November 1994 (1994-11-15), XP002129998</p> <p>&amp; SASAKI, T., ET AL.: "Rice cDNA, partial sequence (S4564_1A)" EMBL ACCESSION NO:D41766, 14 November 1994 (1994-11-14),</p> <p>---</p>	7,10-12, 16,17
Y	<p>SASAKI, T., ET AL.: "Rice cDNA, partial sequence (R2668_1A)" EMBL ACCESSION NO:D24862, 29 November 1993 (1993-11-29), XP002124441 the whole document</p> <p>-&amp; DBEST DATABASE ID:38158, 2 December 1993 (1993-12-02), XP002124442 the whole document</p> <p>---</p> <p style="text-align: center;">-/--</p>	1,2,4-6



## INTERNATIONAL SEARCH REPORT

International Application No

PL., US 99/15871

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 00551 A (MAYO FOUNDATION ;MCDONALD JOHN A (US); SPICER ANDREW P (US); AUGUS) 8 January 1998 (1998-01-08) see page 14 lines 21 and 22, and SEQ ID NO:48 ---	6
P,X	EP 0 875 575 A (NISSHIN SPINNING) 4 November 1998 (1998-11-04) the whole document ---	1,2,4-7, 10-12
P,X	SASAKI, T., ET AL.: "Oryza sativa cDNA, partial sequence (R2825 6A)." EMBL ACCESSION NO:AU031954, 19 October 1998 (1998-10-19), XP002129999 the whole document ---	16,17
P,X	TAYLOR, N.G., ET AL.: "Arabidopsis thaliana cellulose synthase catalytic subunit (IRX3) mRNA, complete cds." EMBL ACCESSION NO:AF088917, 25 May 1999 (1999-05-25), XP002130000 the whole document ---	1,2,6, 16,17
P,X	BLEWITT, M., ET AL.: "BNLGH13827 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (AF027172) cellulose synthase catalytic subunit [Arabidopsis thaliana] gi 4049343 gnl PID e1361041 (AL034567) cellulose synthase catalytic subunit (RSW1) [Arabidopsis thaliana], mRNA sequence." EMBL ACCESSION NO:A1729626, 12 June 1999 (1999-06-12), XP002130001 the whole document ---	1,2,6, 16,17
P,X	BLEWITT, M., ET AL.: "BNLGH15835 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (AF027172) cellulose synthase catalytic subunit [Arabidopsis thaliana] gi 4049343 gnl PID e1361041 (AL034567) cellulose synthase catalytic subunit (RSW1) [Arabidopsis thaliana], mRNA sequence." EMBL ACCESSION NO:A1729981, 12 June 1999 (1999-06-12), XP002130002 the whole document ---	1,2,6, 16,17
P,X	LAOSINCHAI W., ET AL.: "Gossypium hirsutum cellulose synthase catalytic subunit (celA3) mRNA, complete cds." EMBL ACCESSION NO:AF150630, 21 June 1999 (1999-06-21), XP002130003 the whole document ---	1,2,6, 16,17
	---	

-/--

## INTERNATIONAL SEARCH REPORT

International Application No

PC., US 99/15871

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 723 764 A (SINGLETERY GEORGE WILLIAM ET AL) 3 March 1998 (1998-03-03) ---	13
A	AMOR Y ET AL: "EVIDENCE FOR A CYCLIC DIGUANYLIC ACID-DEPENDENT CELLULOSE SYNTHASE IN PLANTS" PLANT CELL,US,AMERICAN SOCIETY OF PLANT PHYSIOLOGISTS, ROCKVILLE, MD, vol. 3, page 989-995 XP002061420 ISSN: 1040-4651 the whole document ---	6,18
A	LI ET AL: "beta-Glucan synthesis in the cotton fiber" PLANT PHYSIOLOGY,US,AMERICAN SOCIETY OF PLANT PHYSIOLOGISTS, ROCKVILLE, MD, vol. 101, no. 4, page 1149-1156 XP002087180 ISSN: 0032-0889 the whole document ---	6,18
A	WO 91 13988 A (UNIV TEXAS) 19 September 1991 (1991-09-19) the whole document -----	1-6, 13-18

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 99/ 15871

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

see additional sheets

1. ☒ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☒ No protest accompanied the payment of additional search fees.

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

**1. Claims: 1-6,13-18 all partially**

Nucleic acid fragments encoding barley cellulose synthase, and corresponding polypeptide as represented by SEQ ID NOs:1 and 2, fragments encoding amino acid sequence 90% identical thereto, transformed hosts containing and methods for obtaining said sequences, methods for altering expression and methods for evaluating inhibitors using said sequences.

**2. Claims: 1-6,13-18 all partially and 7-12 all completely**

Nucleic acid fragments encoding corn cellulose synthase, and corresponding polypeptide as represented by SEQ ID NOs:3-10, fragments encoding amino acid sequence 90% identical thereto, transformed hosts containing and methods for obtaining said sequences, methods for altering expression and methods for evaluating inhibitors using said sequences.

**3. Claims: 1-6,13-18 all partially**

Nucleic acid fragments encoding rice cellulose synthase, and corresponding polypeptide as represented by SEQ ID NOs:11 and 12, fragments encoding amino acid sequence 90% identical thereto, transformed hosts containing and methods for obtaining said sequences, methods for altering expression and methods for evaluating inhibitors using said sequences.

**4. Claims: 1-6,13-18 all partially**

Nucleic acid fragments encoding soybean cellulose synthase, and corresponding polypeptide as represented by SEQ ID NOs:13-18, fragments encoding amino acid sequence 90% identical thereto, transformed hosts containing and methods for obtaining said sequences, methods for altering expression and methods for evaluating inhibitors using said sequences.

**5. Claims: 1-6,13-18 all partially**

Nucleic acid fragments encoding wheat cellulose synthase, and corresponding polypeptide as represented by SEQ ID NOs:19-22, fragments encoding amino acid sequence 90% identical thereto, transformed hosts containing and methods for obtaining said sequences, methods for altering expression and methods for evaluating inhibitors using said sequences.

**6. Claim : 18 partially**

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Method for evaluating a compound for inhibitory activity on cellulose synthase comparing activity of cellulose synthase produced in a transformed host with and without the addition of the compound, not covered by any of the previous groups of claimed inventions 1-5.

# INTERNATIONAL SEARCH REPORT

formation on patent family members

International Application No

PL., US 99/15871

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9800549 A	08-01-1998	AU 3160397 A CA 2259126 A EP 0956353 A	21-01-1998 08-01-1998 17-11-1999
WO 9818949 A	07-05-1998	AU 5092398 A BR 9712457 A EP 0938573 A	22-05-1998 19-10-1999 01-09-1999
WO 9800551 A	08-01-1998	AU 3652297 A	21-01-1998
EP 0875575 A	04-11-1998	JP 10276782 A	20-10-1998
US 5723764 A	03-03-1998	NONE	
WO 9113988 A	19-09-1991	AU 7556991 A	10-10-1991